

526, 572

(12) NACH DEM VEREINBAR ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG(19) Weltorganisation für geistiges Eigentum
Internationales Büro(43) Internationales Veröffentlichungsdatum
18. März 2004 (18.03.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/022586 A2(51) Internationale Patentklassifikation⁷: C07K 14/195(74) Anwälte: BOETERS, Hans, D. usw.; Boeters & Lieck,
Bereiteranger 15, 81541 München (DE).

(21) Internationales Aktenzeichen: PCT/EP2003/009780

(22) Internationales Anmeldedatum:
3. September 2003 (03.09.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 41 152.2 5. September 2002 (05.09.2002) DE(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): GESELLSCHAFT FUER BIOTECHNOL-
OGISCHE FORSCHUNG MBH (GBF) [DE/DE];
Mascheroder Weg 1, 38124 Braunschweig (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): HOEFLE, Gerhard
[DE/DE]; Mascheroder Weg 1, 38124 Braunschweig
(DE). MUELLER, Rolf [DE/DE]; Mascheroder Weg 1,
38124 Braunschweig (DE). SASSE, Florenz [DE/DE];
Mascheroder Weg 1, 38124 Braunschweig (DE). SAND-
MANN, Axel [DE/DE]; Mascheroder Weg 1, 38124
Braunschweig (DE). BLOECKER, Helmut [DE/DE];
Mascheroder Weg 1, 38124 Braunschweig (DE).(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU,
SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,
PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu ver-
öffentlichen nach Erhalt des BerichtsZur Erklärung der Zweibuchstaben-Codes und der anderen Ab-
kürzungen wird auf die Erklärungen ("Guidance Notes on Co-
des and Abbreviations") am Anfang jeder regulären Ausgabe der
PCT-Gazette verwiesen.

Gesellschaft für Biotechnologische Forschung mbH

Tubulysin-Biosynthese-Gene

Tubulysine sind bereits als eine neue, auf das Tubulin-Skelett wirkende Substanzfamilie aus Myxobakterien in Irsee vorgestellt worden; vgl. PCT/EP 97/05095 und DE 100 08 089.8 und die darin angeführte Literatur. Im Gegensatz zu den Epothilonen zeigen diese eine mikrotubuli-abbauende Wirkung sowie die vermehrte Ausbildung von Zentrosomen. Mit einer Cytotoxizität von $IC_{50} = 10 - 500$ pg sind die Tubulysine als potentielle Cytostatika von besonderem Interesse.

Die Tubulysine haben eine cytostatische oder antimitotische Wirkung auf Pilze, humane Tumore oder Krebszelllinien und andere tierische Zellkulturen (vgl. Tabelle). Sie führen in den Zellen zu einem raschen Abbau des Mikrotubuli-Gerüsts. Das Aktinskelett bleibt erhalten. Adhärent wachsende L929-Maus-Zellen vergrößern unter dem Einfluß der Tubulysine ihr Zellvolumen, ohne sich zu teilen, und entwickeln große Zellkerne, die dann in einem apoptischen Vorgang zerfallen.

Wirkunsspektrum

<u>Pilze</u>	<u>Hemmhof [mm]</u>	
	<u>Tubulysin A</u>	<u>Tubulysin B</u>
<i>Aspergillus niger</i>	20	18
<i>Botrytis cineria</i>	23	18
<i>Coprinus cinereus</i>	20	
<i>Pythium debaryanum</i>	20	

Agardiffusionstest: 20 µg pro Testblättchen 6 mm Durchmesser

<u>Humane Krebszellinien</u>	<u>IC₅₀ [ng/ml]</u>		
	<u>Tubulysin A</u>	<u>Tubulysin B</u>	<u>Tubulysin C</u>
KB-3-1 (DSM ACC 158)	0,01	0,02	0,1
K-562 (ATCC CCL 243)	0,1	0,2	1,5
HL-60 (ATCC-CCL 240)	0,04	0,08	0,4

<u>Tierische Zelllinien</u>			
L929, Maus (ATCC CCL1)	0,2	0,4	2
Pt K2, <i>Potorous tri-</i> <i>dactylis</i> (ATCC CCL 56)	0,2	0,2	2

Gemäß einer Ausführungsform betrifft die Erfindung ein ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:

- (i) ss-DNA-Molekül mit einer Sequenz gemäß Figur 1;
- (ii) ss-DNA-Molekül, das mit einem ss-DNA-Molekül gemäß (i) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zu jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von dem ss-DNA-Molekül gemäß (i) hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und
- (iii) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines ss-DNA-Moleküls gemäß (i) oder (ii) komplementär ist.

Ferner betrifft die Erfindung ein ds-DNA-Molekül aus einem erfindungsgemäßen ss-DNA-Molekül und einem dazu komplementären Strang.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung ein ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:

- (i) ss-DNA-Molekül mit einer Sequenz der Positionen 3.308 bis 1 (ORF 16) der Sequenz gemäß Figur 1;
- (ii) ss-DNA-Molekül mit einer Sequenz der Positionen 4706 bis 3453 (ORF 15) der Sequenz gemäß Figur 1;
- (iii) ss-DNA-Molekül mit einer Sequenz der Positionen 5719 bis 7164 (ORF 14) der Sequenz gemäß Figur 1;

- (iv) ss-DNA-Molekül mit einer Sequenz der Positionen 9557 bis 7317 (ORF 13) der Sequenz gemäß Figur 1;
- (v) ss-DNA-Molekül mit einer Sequenz der Positionen 12193 bis 10550 (ORF 12) der Sequenz gemäß Figur 1;
- (vi) ss-DNA-Molekül mit einer Sequenz der Positionen 12841 bis 13881 (ORF 11) der Sequenz gemäß Figur 1;
- (vii) ss-DNA-Molekül mit einer Sequenz der Positionen 14833 bis 13835 (ORF 10) der Sequenz gemäß Figur 1;
- (viii) ss-DNA-Molekül mit einer Sequenz der Positionen 14942 bis 15586 (ORF 9) der Sequenz gemäß Figur 1;
- (ix) ss-DNA-Molekül mit einer Sequenz der Positionen 15847 bis 16983 (ORF 8) der Sequenz gemäß Figur 1;
- (x) ss-DNA-Molekül mit einer Sequenz der Positionen 21154 bis 18809 (ORF 7) der Sequenz gemäß Figur 1;
- (xi) ss-DNA-Molekül mit einer Sequenz der Positionen 22366 bis 23532 (ORF 6) der Sequenz gemäß Figur 1;
- (xii) ss-DNA-Molekül mit einer Sequenz der Positionen 24591 bis 26513 (ORF 5) der Sequenz gemäß Figur 1;
- (xiii) ss-DNA-Molekül mit einer Sequenz der Positionen 26597 bis 27517 (ORF 4) der Sequenz gemäß Figur 1;

- (xiv) ss-DNA-Molekül mit einer Sequenz der Positionen 29858 bis 30400 (ORF 3) der Sequenz gemäß Figur 1;
- (xv) ss-DNA-Molekül mit einer Sequenz der Positionen 31220 bis 32392 (TubA) der Sequenz gemäß Figur 1;
- (xvi) ss-DNA-Molekül mit einer Sequenz der Positionen 33056 bis 32397 (ORF 2) der Sequenz gemäß Figur 1;
- (xvii) ss-DNA-Molekül mit einer Sequenz der Positionen 34195 bis 33074 (TubZ) der Sequenz gemäß Figur 1;
- (xviii) ss-DNA-Molekül mit einer Sequenz der Positionen 35422 bis 34205 (ORF 1) der Sequenz gemäß Figur 1;
- (xix) ss-DNA-Molekül mit einer Sequenz der Positionen 35522 bis 40147 (TubB) der Sequenz gemäß Figur 1;
- (xx) ss-DNA-Molekül mit einer Sequenz der Positionen 40144 bis 48021 (TubC) der Sequenz gemäß Figur 1;
- (xxi) ss-DNA-Molekül mit einer Sequenz der Positionen 48011 bis 58558 (TubD) der Sequenz gemäß Figur 1;
- (xxii) ss-DNA-Molekül mit einer Sequenz der Positionen 58551 bis 62096 (TubE) der Sequenz gemäß Figur 1;
- (xxiii) ss-DNA-Molekül mit einer Sequenz der Positionen 62103 bis 70616 (TubF) der Sequenz gemäß Figur 1;
- (xxiv) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x),

(xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii) oder (xxiii) unter stringenten Bedingungen hybridisierbar ist und insbesondere dieselbe Anzahl von Basen aufweist; und

(xxv) ss-DNA-Molekül, das mit einem ss-DNA-Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii) oder (xxiii) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zu jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von diesem ss-DNA-Molekül hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und

(xxvi) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines Moleküls gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv) oder (xxv) komplementär ist.

Ferner betrifft die Erfindung ein ds-DNA-Molekül aus einem derartigen erfindungsgemäßen ss-DNA-Molekül und einem dazu komplementären Strang.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung ein ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:

- (i) ss-DNA-Molekül mit einer Sequenz der Positionen 35747 bis 36769 (Domäne C des tubB-Gens) der Sequenz gemäß Figur 1;
- (ii) ss-DNA-Molekül mit einer Sequenz der Positionen 37184 bis 39817 (Domäne A des tubB-Gens) der Sequenz gemäß Figur 1;
- (iii) ss-DNA-Molekül mit einer Sequenz der Positionen 38369 bis 39730 (Domäne NMT des tubB-Gens) der Sequenz gemäß Figur 1;
- (iv) ss-DNA-Molekül mit einer Sequenz der Positionen 39818 bis 40069 (Domäne PCP des tubB-Gens) der Sequenz gemäß Figur 1;
- (v) ss-DNA-Molekül mit einer Sequenz der Positionen 40372 bis 41397 (Domäne C des tubC-Gens) der Sequenz gemäß Figur 1;
- (vi) ss-DNA-Molekül mit einer Sequenz der Positionen 41824 bis 43215 (Domäne A des tubC-Gens) der Sequenz gemäß Figur 1;
- (vii) ss-DNA-Molekül mit einer Sequenz der Positionen 43216 bis 43461 (Domäne PCP des tubC-Gens) der Sequenz gemäß Figur 1;
- (viii) ss-DNA-Molekül mit einer Sequenz der Positionen 43552 bis 44574 (Domäne C des tubC-Gens) der Sequenz gemäß Figur 1;

- (ix) ss-DNA-Molekül mit einer Sequenz der Positionen 44980 bis 47631 (Domäne A des tubC-Gens) der Sequenz gemäß Figur 1;
- (x) ss-DNA-Molekül mit einer Sequenz der Positionen 46153 bis 47547 (Domäne NMT des tubC-Gens) der Sequenz gemäß Figur 1;
- (xi) ss-DNA-Molekül mit einer Sequenz der Positionen 47632 bis 47868 (Domäne PCP des tubC-Gens) der Sequenz gemäß Figur 1;
- (xii) ss-DNA-Molekül mit einer Sequenz der Positionen 48011 bis 49321 (Domäne KS des tubD-Gens) der Sequenz gemäß Figur 1;
- (xiii) ss-DNA-Molekül mit einer Sequenz der Positionen 49622 bis 50584 (Domäne AT des tubD-Gens) der Sequenz gemäß Figur 1;
- (xiv) ss-DNA-Molekül mit einer Sequenz der Positionen 51473 bis 52309 (Domäne KR des tubD-Gens) der Sequenz gemäß Figur 1;
- (xv) ss-DNA-Molekül mit einer Sequenz der Positionen 53066 bis 53980 (Domäne ER des tubD-Gens) der Sequenz gemäß Figur 1;
- (xvi) ss-DNA-Molekül mit einer Sequenz der Positionen 54158 bis 54460 (Domäne ACP des tubD-Gens) der Sequenz gemäß Figur 1;

- (xvii) ss-DNA-Molekül mit einer Sequenz der Positionen 54461 bis 55870 (Domäne HC des tubD-Gens) der Sequenz gemäß Figur 1;
- (xviii) ss-DNA-Molekül mit einer Sequenz der Positionen 56000 bis 57412 (Domäne A des tubD-Gens) der Sequenz gemäß Figur 1;
- (xix) ss-DNA-Molekül mit einer Sequenz der Positionen 57413 bis 57643 (Domäne PCP des tubD-Gens) der Sequenz gemäß Figur 1;
- (xx) ss-DNA-Molekül mit einer Sequenz der Positionen 58689 bis 59714 (Domäne C des tubE-Gens) der Sequenz gemäß Figur 1;
- (xxi) ss-DNA-Molekül mit einer Sequenz der Positionen 60156 bis 61697 (Domäne A des tubE-Gens) der Sequenz gemäß Figur 1;
- (xxii) ss-DNA-Molekül mit einer Sequenz der Positionen 61698 bis 61967 (Domäne PCP des tubE-Gens) der Sequenz gemäß Figur 1;
- (xxiii) ss-DNA-Molekül mit einer Sequenz der Positionen 62127 bis 63320 (Domäne KS des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxiv) ss-DNA-Molekül mit einer Sequenz der Positionen 63711 bis 64676 (Domäne AT des tubF-Gens) der Sequenz gemäß Figur 1;

- (xxv) ss-DNA-Molekül mit einer Sequenz der Positionen 64959 bis 65882 (Domäne KR des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxvi) ss-DNA-Molekül mit einer Sequenz der Positionen 65985 bis 67061 (Domäne CMT des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxvii) ss-DNA-Molekül mit einer Sequenz der Positionen 67242 bis 67829 (Domäne DH des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxviii) ss-DNA-Molekül mit einer Sequenz der Positionen 68247 bis 69128 (Domäne ER des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxix) ss-DNA-Molekül mit einer Sequenz der Positionen 69360 bis 69605 (Domäne PCP des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxx) ss-DNA-Molekül mit einer Sequenz der Positionen 69759 bis 70586 (Domäne TE des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxxi) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix) oder (xxx) unter stringenten Bedingungen hybridisierbar ist und insbesondere dieselbe Anzahl von Basen aufweist;

(xxxii) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xiv), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix) oder (xxx) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zur jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von diesem ss-DNA-Molekül hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und

(xxxiii) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines Moleküls gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xiv), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix), (xxx), (xxxi) oder (xxxii) komplementär ist.

Ferner betrifft die Erfindung ein ds-DNA Molekül aus einem derartigen ss-DNA-Molekül und einem dazu komplementärem Strang.

Ferner betrifft die Erfindung Varianten oder Mutanten, die aus einer Substitution, Insertion oder Deletion von Nucleotiden oder einer Inversion von Nucleotid-Segmenten eines erfindungsgemäßen ss-DNA-Moleküls oder eines erfindungsgemäßen ds-DNA-Moleküls resultieren, wobei diese Varianten und Mutanten Enzym-Varianten oder Enzym-Mutanten für die Produktion von Sekundärstoff(en) mit den eingangs geschilderten und für Tubulysine charakteristischen Eigenschaften codieren, insbesondere mit der cytostatischen Wirkung. Der Fachmann ist mit Massenscreening vertraut.

Ferner betrifft die Erfindung RNA

- (a) mit einer Sequenz entsprechend der eines erfindungsgemäßen ss-DNA-Moleküls oder
 - (b) mit einer Sequenz einer RNA gemäß (a), aber in Gegenrichtung (anti-sense), oder
 - (c) mit einer Sequenz einer RNA gemäß (a) oder (b) und mit einem dazu komplementären Strang,
- jeweils gegebenenfalls als Element eines rekombinanten Vektors.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung einen rekombinanten Vektor, insbesondere Expressionsvektor mit einem erfindungsgemäßen DNA-Molekül.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung eine Zelle, insbesondere zur Expression, in die ein erfindungsgemäßes DNA-Molekül oder ein erfindungsgemäßer Vektor integriert ist.

Die erfindungsgemäße Zelle kann sich von kultivierbaren Bakterien wie Myxobakterien wie Angiococcus, insbesondere A. disciformis, Archangium, insbesondere A. gephyra, Escherichia coli, Pseudomonaden oder Actinomyceten herleiten.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung eine Verwendung eines erfindungsgemäßen Vektors für die Transformation von Zellen oder Organismen zur transienten oder permanenten Expression eines oder mehrerer Proteine (Expressionsprodukt(e)), das (die) durch eine DNA (ssDNA oder dsDNA) des Vektors codiert wird (werden).

Gemäß einer weiteren Ausführungsform betrifft die Erfindung eine Verwendung einer erfindungsgemäßen Zelle zur enzymatischen

Biosynthese, Mutasynthese oder Partial-Synthese eines Tubulysins, insbesondere Tubulysin A, B, C, D, E und/oder F.

Gemäß einer weiteren Ausführungsform betrifft die Erfindung ein Expressionsprodukt eines erfindungsgemäßen DNA-Moleküls oder eines erfindungsgemäßen Vektors oder einer erfindungsgemäßen Zelle.

Die vorliegende Erfindung betrifft insbesondere ein Polynucleotid, das eine Sequenz wie in SEQ ID NO: 1, 18, 33 oder 36 definiert, oder ein Fragment davon enthält. SEQ ID NO: 1 und 18 beschreiben den (+)- bzw. den (-)-Strang des Tubulysin Biosynthese-Clusters von *Angiococcus disciformis*. SEQ ID NO: 33 ist eine Sequenz, die mehrere sich überlappende Gene des Clusters enthält. SEQ ID NO: 36 beschreibt eine Mutante von *Angiococcus disciformis*. Überraschenderweise wurde gefunden, dass diese Mutante ein Vielfaches der Tubulysin D-Produktion des Wildtyps zeigte. Die Tubulysin-Überexpression ist auf die Gesamtwirkung aller Tubulysin-Derivate bezogen sogar noch höher als die von Tubulysin D, was in keinsterweise zu erwarten war. Die Gene der SEQ ID NO: 36 sind offenbar an der negativen Regulation der Tubulysin Expression beteiligt. Diese Mutante ist auf Grund der erhöhten Expression aller Tubulysine besonders geeignet zur Produktion der erfindungsgemäßen Polypeptide. Antikörper gegen die Wildtyp-Expressionsprodukte dieser Sequenz können eingesetzt werden, um deren negativen Einfluss auf die Tubulysin Produktion auch in anderen Stämmen zu minimieren. Einen ähnlichen Effekt haben auch antisense-RNA bzw RNAi-Techniken, die mit der Wildtyp Sequenz der negativen Regulatorgene interagieren.

Die Fragmente des Polynucleotid können jede beliebige Teilsequenz und Länge haben, bevorzugt sind jedoch solche Fragmente, die Proteine codieren. Die Polynucleotide der vorliegenden Erfindung schließen ebenfalls ein, sind jedoch nicht beschränkt auf, ein Polynucleotid, das an den komplementären Strang der offenbaren Nucleotidsequenzen unter moderat stringenten oder stringenten Bedingungen hybridisiert; ein Polynucleotid, das eine Allel-Variante irgendeines oben beschriebenen Polynucleotids ist; ein Polynucleotid, das ein Spezies-Homolog irgendwelcher der hier offenbaren Proteine codiert; oder ein Polynucleotid, das ein Polypeptid codiert, das eine zusätzliche spezifische Domäne oder eine Trunkierung bzw. Verkürzung der offenbaren Proteine aufweist.

Der Begriff "CDS" bezeichnet eine Sequenz von Nucleotiden, die der Sequenz von Aminosäuren in einem Protein entspricht, also die Aminosäuren-codierenden Sequenzbereiche, inklusive des jeweiligen Start- und Stoppcodons.

In einer bevorzugten Ausführungsform ist das erfindungsgemäße Polynucleotid ein Fragment, das eine im Sequenzprotokoll definierte CDS ist.

Die vorliegende Erfindung betrifft weiterhin einen Vektor, der ein Polynucleotid wie oben beschrieben enthält. Vektoren für verschiedene Zwecke sind im Stand der Technik bekannt, ebenso die Techniken zur Subklonierung von Polynucleotiden in solche Vektoren. Diese werden beispielsweise in der neuen Ausgabe von Molecular Cloning: A Laboratory Manual, (Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press); DNA Cloning, Volumes I and II (D. N. Glover ed., 1985); Gene Transfer Vectors for Mammalian Cells

(Miller & Calos, eds.); Current Protocols in Molecular Biology and Short Protocols in Molecular Biology, 3rd Edition (F. M. Ausubel et al., eds.); Recombinant DNA Methodology (R. Wu ed., Academic Press) oder "A Practical Guide To Molecular Cloning" beschrieben. Beispiele für Vektoren finden sich unter anderen in Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory).

Bei dem Vektor handelt es sich vorzugsweise um einen Expressionsvektor, also im Allgemeinen ein Plasmid, einen Phagen, ein Virus oder einen Vektor zum Exprimieren eines Polypeptids aus einer DNA (RNA) Sequenz. Ein Expressionsvektor kann eine Transkriptionseinheit umfassen, die eine Anordnung des Folgenden aufweist: (1) ein genetisches Element oder Elemente mit einer regulatorischen Rolle in der Genexpression, beispielsweise Promotoren oder Enhancer, (2) eine Struktursequenz oder codierende Sequenz, die in mRNA transkribiert und in ein Protein translatiert wird und (3) geeignete Transkriptionsstart- und -terminationssequenzen. Struktureinheiten, die zur Verwendung in Hefen oder eukaryontischen Expressionssystemen vorgesehen sind, schließen vorzugsweise eine Leadersequenz ein, die die extrazelluläre Sekretion eines translatierten Proteins durch einen Wirt ermöglicht. Es kann alternativ, wenn ein rekombinantes Protein ohne eine Leader- oder Transportsequenz exprimiert wird, einen N-terminalen Methionin-Rest einschließen. Dieser Rest kann, muss aber nicht anschließend von dem exprimierten rekombinanten Protein abgespalten werden, um das Endprodukt zu erhalten.

Des Weiteren betrifft die vorliegende Erfindung eine Zelle, die einen solchen Vektor enthält. Der Vektor kann über die bekannten Techniken wie beispielsweise Transfektion, Elektroporation,

Lipofektion usw. in die Zelle eingebracht werden. Bei viralen Vektoren ist auch eine Infektion möglich. Bei den Zellen kann es sich um eukaryotische als auch prokaryotische Zellen handeln. Auch die Methoden zur Selektion und Propagierung der Zellen, die den Vektor enthalten sind dem Fachmann bekannt. Beispiele für die Kultivierung von Zellen tierischen Ursprungs sind u.a. in Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987) zu finden.

Eine weitere Ausführungsform betrifft ein Polypeptid, das wenigstens eine Sequenz wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder ein Fragment und/oder Derivat davon enthält. Das Polypeptid kann sowohl durch Expression eines Polynucleotids als auch durch chemische Synthese bereitgestellt werden.

Die Aminosäuresequenzen der vorliegenden Erfindung umfassen ebenfalls alle Sequenzen, die sich von den hier offenbarten Sequenzen durch Aminosäureinsertionen, -deletionen und -substitutionen unterscheiden.

Aminosäure- „Substitutionen“ sind vorzugsweise das Ergebnis des Ersetzens einer Aminosäure durch eine andere Aminosäure mit ähnlichen strukturellen und/oder chemischen Eigenschaften, das heißt konservative Aminosäure-Austausche. Aminosäure-Substitutionen können auf der Grundlage einer Ähnlichkeit in der Polarität, Ladung, Löslichkeit, Hydrophobie, Hydrophilie und/oder der amphipatischen Natur der einbezogenen Reste vorgenommen werden. Beispielsweise schließen unpolare (hydrophobe) Aminosäuren Alanin, Leucin, Isoleucin, Valin, Prolin, Phenylalanin, Tryptophan und Methionin ein; polare neutrale Aminosäuren schließen Glycin, Serin, Threonin, Cystein,

Tyrosin, Asparagin und Glutamin ein; positiv geladene (basische) Aminosäuren schließen Arginin, Lysin und Histidin ein; und negativ geladene (saure) Aminosäuren schließen Asparaginsäure und Glutaminsäure ein.

"Insertionen" oder "Deletionen" bewegen sich typischerweise im Bereich von 1-3 Aminosäuren. Die erlaubte Variation kann experimentell bestimmt werden, indem systematisch Insertionen, Deletionen oder Substitutionen von Aminosäuren in einem Polypeptidmolekül unter Verwendung von DNA-Rekombinationstechniken vorgenommen und die sich ergebenden rekombinanten Varianten bezüglich ihrer Aktivität untersucht werden. Dazu ist für den Fachmann nicht mehr als die Durchführung von Routineexperimenten erforderlich.

Beispielsweise kann das Polypeptid auch in Form eines chimären Polypeptids, codiert durch ein Fusionsgens, vorliegen, dass wenigstens eine weitere Sequenz enthält. Diese zusätzliche Sequenz kann z.B. dazu dienen, die Aufreinigung des Expressionsprodukts zu erleichtern oder dem Expressionsprodukt eine zusätzlich Funktion zu verleihen.

Beispiele für zusätzliche Sequenzen, die die Aufreinigung erleichtern sind sogenannte Tags, die dem Fachmann bekannt sind, wie z.B. das his-tag.

Die vorliegende Erfindung betrifft außerdem die Verwendung wenigstens eines Polynucleotids wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens eines Fragments davon und/oder wenigstens eines Polypeptids wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens eines Fragments davon zur Herstellung einer pharmazeutischen Zusammensetzung zur Behandlung von

unerwünschtem Zellwachstums oder unerwünschter Zellvermehrung in einem Individuum. Die Zusammensetzung kann beispielsweise einen geeigneten Vektor enthalten, zusammen mit Hilfsfaktoren, die die Expression eines Tubulysins, vorzugsweise in den unerwünschten Zellen, ermöglichen und dadurch ein weiteres Wachstum bzw. eine weitere Vermehrung dieser Zellen verhindern. Die Zusammensetzung kann auch erfindungsgemäße Zellen enthalten, die mit einem Vektor, beispielsweise einen Tubulysin-exprimierenden Vektor, transfiziert wurden.

In einer bevorzugten Ausführungsform ist das unerwünschte Zellwachstum oder die unerwünschte Zellvermehrung ein Tumor. Der Tumor kann sowohl eine gutartige als auch eine bösartige Geschwulst sein.

In einer weiteren Ausführungsform ist das unerwünschte Zellwachstum eine pathogene Infektion. Das Pathogen kann hierbei sowohl einzellig als auch mehrzellig sein. Dies schließt auch Infektionen mit Pilzen, wie beispielsweise Candida oder Aspergillus, und Infektionen mit Parasiten, wie beispielsweise Trypanosomen oder Schistosomen ein. In einer bevorzugten Ausführungsform der Verwendung ist die pathogene Infektion eine Mykose, Malaria oder eine parasitäre Erkrankung.

Die Erfindung betrifft ferner eine pharmazeutische Zusammensetzung, enthaltend wenigstens ein Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens ein Fragment davon und/oder wenigstens ein Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens ein Fragment davon. Die Zusammensetzungen enthalten eine therapeutisch wirksame Menge bzw. Dosis des jeweiligen Wirk- bzw. Inhaltsstoffes. Eine therapeutisch

wirksame Dosis betrifft diejenige Menge der Verbindung, die ausreicht, um eine Linderung der Symptome, beispielsweise eine Behandlung, Heilung, Prävention oder Linderung derartiger Zustände, insbesondere die Hemmung oder Verhinderung von unerwünschtem Zellwachstum und Zellvermehrung in einem Patienten zu ergeben. Geeignete Verabreichungswege schließen beispielsweise eine parenterale Verabreichung, einschließlich intramuskulärer und subkutaner Injektionen, ebenso wie intrathekaler, direkt intraventrikulärer, intravenöser, intraperitonealer Injektionen ein.

In einer weiteren Ausführungsform enthält diese pharmazeutische Zusammensetzung wenigstens einen pharmazeutisch verträglichen Trägerstoff. Eine solche Zusammensetzung kann weiterhin (zusätzlich zum Inhaltsstoff und dem Träger) Verdünnungsmittel, Füllmittel, Salze, Puffer, Stabilisatoren, Lösungsvermittler und andere Materialien enthalten, die im Stand der Technik wohl bekannt sind. Der Begriff "pharmazeutisch verträglich" bedeutet ein nicht-toxisches Material, das die Wirksamkeit der biologischen Aktivität des aktiven Inhaltsstoff (der aktiven Inhaltsstoffe) nicht stört. Die Eigenschaften des Trägers hängen vom Verabreichungsweg ab. Die therapeutische Zusammensetzung kann weiterhin weitere Mittel bzw. Wirkstoffe enthalten, die die Aktivität bzw. Wirksamkeit oder Anwendung bei Behandlung verbessern bzw. erleichtern. Solche zusätzlichen Faktoren und/oder Mittel können in der therapeutischen Zusammensetzung enthalten sein, um eine synergistische Wirkung zu erzeugen oder um Nebenwirkungen zu minimieren.

Techniken zur Formulierung, Zubereitung und Verabreichung der Verbindungen der vorliegenden Erfindung sind in "Remington's

Pharmaceutical Sciences", Mack Publishing Co., Easton, PA, neueste Ausgabe, zu finden.

Die vorliegende Erfindung betrifft außerdem ein Verfahren zur Herstellung von Tubulysinen und Tubulysin-Biosynthese Proteinen, umfassend die Schritte:

- (a) exprimieren von wenigstens einem Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens einem Fragment davon und/oder wenigstens einem Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens einem Fragment davon, und
- (b) aufreinigen der Expressionsprodukte.

Methoden zur Expression von Proteinen sind dem Fachmann bekannt und können der einschlägigen Literatur entnommen werden, beispielweise aus Methods In Enzymology, Vols. 154 and 155 (Wu et al. eds.) oder Recombinant DNA Methodology (R. Wu ed., Academic Press). Für das Aufreinigen von Expressionsprodukten sind dem Fachmann eine Vielzahl von Methoden bekannt. Neben chromatographischen Methoden wie beispielsweise Affinitätschromatographie oder HPLC können auch immunologische Verfahren wie beispielsweise immobilisierte Antikörper gegen ein Epitop auf dem Expressionsprodukt, z.B. einem His-tag, zum Aufreinigen der Produkte verwendet werden.

In einer bevorzugten Ausführungsform erfolgt die Expression in prokaryotischen oder eukaryotischen Zellen und/oder durch in vitro Expression. Die Expression von Polypeptiden in prokaryotischen oder eukaryotischen Zellen ist ein häufig angewandtes Verfahren und wird im Allgemeinen mittels eines Expressionsvektors, wie oben beschrieben, erreicht. Für die in vitro Expression sind ebenfalls bereits Vektoren beschrieben.

Diese und die notwendigen Faktoren sind kommerziell in Form von Kits erhältlich, beispielsweise von BioRad, Stratagene, Invitrogen und Clontech.

Desweiteren stellt die Erfindung ein Verfahren zum Auffinden von Genen, die an der Biosynthese von Tubulysinen beteiligt sind, bereit. Das Verfahren umfasst die folgenden Schritte:

- (a) hybridisieren von wenigstens einem Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens einem Fragment davon mit DNA, RNA, und/oder cDNA einer Spezies, die nicht mit *Angiococcus disciformis* identisch ist, und
- (b) isolieren und charakterisieren der hybridisierten DNA, RNA, und/oder cDNA.

Die Hybridisierung kann unter unterschiedlich stringenten Bedingungen durchgeführt werden.

Die Stringenz der Hybridisierung, wie hierin verwendet, betrifft Bedingungen, unter denen Polynucleotid-Doppelstränge stabil sind. Wie dem Fachmann bekannt ist, ist die Stabilität eines Doppelstranges eine Funktion der Natriumionenkonzentration und der Temperatur (siehe beispielsweise Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd Ed. (Cold Spring Harbor Laboratory, (1989)). Die Stringenzniveaus, die zur Hybridisierung verwendet werden, können vom Fachmann leicht abgewandelt werden.

Der Begriff „schwach stringente Hybridisierung“ bezeichnet Bedingungen, die einer Hybridisierung in 10% Formamid, 5 x Denharts Lösung, 6 x SSPE, 0,2 % SDS bei 42°C, gefolgt von Waschen in 1 x SSPE, 0,2 % SDS bei 50°C äquivalent sind.

Denhart's Lösung und SSPE sind dem Fachmann genauso wie andere geeignete Hybridisierungspuffer wohl bekannt.

Eine „moderat stringente Hybridisierung“ bedeutet Bedingungen, die es der DNA erlauben, an eine komplementäre Nucleinsäure zu binden, die ungefähr 60 % Identität, vorzugsweise ungefähr 75 % Identität, besonders bevorzugt ungefähr 85 % Identität zu dieser DNA aufweist; wobei eine Identität von mehr als ungefähr 90 % zu dieser DNA besonders bevorzugt wird. Moderat stringente Bedingungen sind vorzugsweise Bedingungen, die eine Hybridisierung in 50 % Formamid, 5 x Denharts Lösung, 5 x SSPE, 0,2 % SDS bei 42°C gefolgt von Waschen in 0,2 x SSPE, 0,2 % SDS bei 65°C äquivalent sind.

Hochstringente Hybridisierung bedeutet Bedingungen, die die Hybridisierung nur von solchen Nucleinsäuresequenzen ermöglichen, die in 0,018 M NaCl bei 65°C stabile Doppelstränge bilden (d.h., wenn ein Doppelstrang in 0,018 M NaCl bei 65°C nicht stabil ist, ist er unter den hier beschriebenen/definierten hochstringenten Bedingungen nicht stabil).

Weiterhin können Nucleinsäure-Hybridisierungstechniken verwendet werden, um eine Nucleinsäure zu identifizieren und zu gewinnen, die von der vorliegenden Erfindung umfasst ist. Kurz gesagt kann jede Nucleinsäure mit einer gewissen Homologie zu einer in dieser Erfindung offenbarten Sequenz oder einem Fragment davon, als Sonde zur Identifizierung einer ähnlichen Nucleinsäure durch Hybridisierung unter moderat stringenten bis hochstringenten Bedingungen verwendet werden. Solche ähnlichen Nucleinsäuren können dann isoliert, sequenziert und analysiert werden, um zu bestimmen, ob sie von der vorliegenden Erfindung umfasst werden.

Die vorliegende Erfindung stellt außerdem einen Kit zur Herstellung von Tubulysinen, enthaltend:

- (a) wenigstens ein Polynucleotid, enthaltend eine Sequenz wie in SEQ ID NO: 1, 18, 33 oder 36 definiert oder ein Fragment davon und/oder wenigstens einen Vektor der ein solches Polynucleotid enthält oder
- (b) geeignete Medien und Puffer zur Vermehrung von Zellen, die eine Expression des Polynucleotids und/oder Vektors erlauben, und
- (c) geeignete Mittel zum Aufreinigen des/der Expressionsprodukte(s)

bereit.

Auf Grund ihrer Wirkung auf das Tubulin-Skelett und ihrer Cytotoxizität, insbesondere bei Pilzen, eignen sich Tubuline auch als Desinfektionsmittel, das eine Kontamination mit Tubulin-enthaltenden Zellen abbauen bzw. verhindern kann. Die Erfindung betrifft deshalb auch die Verwendung einer Zusammensetzung enthaltend wenigstens ein Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens ein biologisch aktives Fragment oder Derivat davon als Desinfektionsmittel. Neben dem oben definierten Polypeptid können im Desinfektionsmittel auch andere desinfizierend wirkende Stoffe enthalten sein, solange sie die Wirkung des erfindungsgemäßen Polypeptids nicht inhibieren. Außerdem kann das Desinfektionsmittel weitere Hilfsstoffe wie beispielsweise Puffer, Wasser, Farbstoffe, Duftstoffe, Stabilisatoren, Trägerstoffe usw. enthalten. In einer bevorzugten Ausführungsform ist die Zusammensetzung flüssig oder pulverförmig.

Entsprechend betrifft die Erfindung auch Desinfektionsmittel wie oben definiert.

Soweit keine anderen Definitionen angegeben sind, weisen alle technischen und wissenschaftlichen Begriffe, die hierin verwendet werden, dieselbe Bedeutung auf, wie sie üblicherweise vom Fachmann auf dem Gebiet, an den sich diese Erfindung wendet, verstanden werden. Alle Veröffentlichungen, Patentanmeldungen, Patente und anderen hierin erwähnten Referenzen sind durch Bezugnahme in ihrer Gesamtheit mit aufgenommen. Im Falle eines Konfliktes wird jedoch die vorliegende Beschreibung einschließlich der Definitionen entscheiden. Zusätzlich sind die Materialien, Methoden und Beispiele lediglich veranschaulichend und sollen nicht als einschränkend aufgefasst werden.

1. Identifizierung des Tubulysin-Biosyntheseclusters in *Angiococcus disciformis* An d48 durch mariner Transposon-Mutagenese mittels pMycoMar

Die Identifizierung des Tubulysin-Biosyntheseclusters wurde durch die Erstellung einer Transposon-Mutantenbank aus *Angiococcus disciformis* An d48 mittels pMycoMar durchgeführt.

Rubin & Mekalanos (Proc. Natl. Acad. Sci. USA, 96 (1999), 1645 - 1650) entwickelten aus dem mariner Element *Himar1* das Plasmid pMycoMar, welches ein einfaches Transposonssystem darstellt, das effizient Bakterien *in vivo* infizieren und Insertions-Mutanten generieren kann. Dieses Plasmid enthält das Mini-Transposon *magellan4*, bei dem das Tn5 Kanamycin-Resistenzgen und der *oriR6K* von den inverted repeats des *Himar1* flankiert sind. Zusätzlich ist die *Himar1* Transposase unter der transkriptionalen Kontrolle des T6 Promotors in das mycobakterielle Temperatur-sensitive Replicon pPR23 kloniert worden. Das pMycoMar codiert ebenfalls ein Gentamycin-Resistenzgen.

Das *Himar1* zeichnet sich bei der Transposition durch eine TA Dinucleotid-Erkennungssequenz aus. Daher kann es zufällig in ein Wirts-Genom integrieren und statistisch gesehen alle aktiven Gene durch eine Insertions-Mutation ausschalten. Aufgrund dieser Tatsache sollte eine Mutantenbank aus An d48 generiert und das Tubulysin-Biosynthesecluster durch eine Knockout-Mutante identifiziert werden.

Alternativ kann man auch von *Archangium gephyra* DSM 11092 ausgehen und nach einem Protokoll von Biozym Diagnostic (Oldendorf, DE; Katalog TSM99K2; pEZ::TN<KAN-2> Tnp Transposome-Kit) arbeiten.

1.1 Generierung der Mutantenbank

Für eine Elektrottransformation von *A. disciformis* An d48 wurden zwei verschiedene Protokolle verwendet. Diese Protokolle wurden für die Myxobakterien *Stigmatella aurantiaca* (Stamm & Plaga, Arch. Microbiol., 172 (1995), 483 - 494) und *Myxococcus xanthus* (Kashefi & Hartzell, Mol. Microbiology, 15 (1995) 483 - 494) etabliert. Die beiden Methoden zeigten in der Transformationseffizienz des *A. disciformis* An d48 keinen Unterschied, so dass die Elektrottransformation zur Erstellung der Transposonbank nach dem Protokoll für *Stigmatella aurantiaca* durchgeführt wurde. Im folgenden sind die beiden Protokolle aufgeführt.

1.1.1. Elektrottransformation von *Angiococcus disciformis* An d48 nach *Stigmatella aurantiaca*-Protokoll

Eine in 50 ml Tryptonmedium (10 g Trypton; 2 g MgSO_4 , 0,1 % Vitamin B12 [10 ng/ml]; 0,2 % Glucose auf 1 l Medium; pH 7,2) gewachsene *A. disciformis* Kultur wird bis $2 \cdot 10^8$ Zellen / ml bei 30°C kultiviert. Ausgehend von einer Generationszeit von 6 Stunden, wurde diese Kultur am Vortag so angeimpft, dass rechnerisch diese Zelldichte erreicht wird. Die Kultur wird bei 20°C abzentrifugiert (20min; 4000rpm) und die Zellen im gleichen Volumen Waschpuffer (5 mM HEPES/NaOH, 0,5 mM CaCl_2 , pH 7,2) resuspendiert. Nach erneuter Zentrifugation wird in 25 ml Puffer resuspendiert und erneut abzentrifugiert. Vor diesem Zentrifugationsschritt wird die absolute Zellzahl innerhalb der 25 ml bestimmt, so dass rechnerisch $1 \cdot 10^9$ Zellen/ 40 μl resuspendiert werden.

Elektroporationsbedingungen:

1-3 µg DNA und 40 µl Zellsuspension werden gemischt und in eine auf Eis gekühlte Elektroporationsküvette (0,1 cm) gegeben. Die Elektroporation wird bei 200 Ω, 25 mF und 0,85 kV / cm durchgeführt.

Direkt nach der Elektroporation wird 1 ml Tryptonmedium zugegeben. Nach Transfer in 50 ml Tryptonmedium werden die Zellen für eine phänotypische Expression 6 h bei 30°C geschüttelt. Danach wird die Kultur abzentrifugiert (20 min, 4000 rpm, 20°C) und in 1 ml Trypton resuspendiert. Ausgehend von einer 100 %-igen Überlebensrate der Zellen wird eine Verdünnungsreihe erstellt und $1 \cdot 10^8$ - $1 \cdot 10^4$ Zellen werden mit 3 ml Trypton-Softagar auf Kanamycin haltigen (50 µg / ml) Tryptonplatten plattiert. Diese Platten werden bei 30°C inkubiert und nach 5 - 8 Tagen sind die ersten Klone zu sehen.

1.1.2. Elektrotransformation von *Angiococcus disciformis* An d48 nach *Myxococcus xanthus*-Protokoll

Die Wachstumsbedingungen der Vor- und Hauptkultur sowie Zentrifugationen und das abschließende Einengen der Zellzahl wurden genauso durchgeführt wie unter 1.1.1. angegeben ist. Dies ist abweichend vom Standardprotokoll für *Myxococcus xanthus* optimiert worden.

Elektroporationsbedingungen:

1-3 µg DNA und 40 µl Zellsuspension werden gemischt und in eine auf Eis gekühlte Elektroporationsküvette (0,1 cm) gegeben. Die Elektroporation wird bei 400 Ω, 25 µF und 0,65kV / cm durchgeführt.

Direkt nach der Elektroporation wird 1 ml Tryptonmedium zugegeben und in einem 1,5 ml Eppendorf-Reagenzgefäß für 6 h bei 30°C geschüttelt. Ausgehend von einer 100%-igen Überlebensrate der Zellen wird eine Verdünnungsreihe erstellt und $1 \cdot 10^8$ - $1 \cdot 10^4$ Zellen mit 3 ml Trypton-Softagar auf Kanamycin haltigen (50 µg / ml) Tryptonplatten plattiert. Diese Platten werden bei 30°C inkubiert und nach 5 - 8 Tagen sind die ersten Klone / Mutanten zu sehen, die mittels Impföse gepickt wurden.

1.2 Kultivierung der Transposonmutanten

Die Mutanten wurden in 96'er Mikrotiter-Platten in 200 µl M7-Medium (5 g Probion; 1 g CaCl_2 ; 1 g MgSO_4 ; 1g Hefeextrakt; 5 g Stärke; 10 g HEPES; 0,1 % Vitamin B12 [10 ng / ml] auf 1 l Medium; pH7,4) bei 32°C inkubiert und nach 10 Tagen eine Kopie der gesamten Bank erstellt. Dafür wurden 50 µl Kultur jeder Mutante in neue Mikrotiter-Platten mit 100 µl M7-Medium transferiert. Nach weiteren sieben Tagen Inkubation wurde eine Kopie bei -80°C als Dauerkulturen eingefroren. Die verbleibende Kopie der Bank wurde extrahiert, der Extrakt mittels Toxizitätstest auf generierte Tubulysin Knockout-Mutanten untersucht.

Bei der Identifizierung von Mutanten, die in dieser Analytik Veränderungen zum Wildtyp zeigten (keine Zellkernfragmentierung), wurden diese aus der Dauerkultur rekultiviert. Zur Kontrolle der erzielten Ergebnisse sollten 50 ml M7-Medium Großkulturen der entsprechenden Mutanten erneut getestet werden. Bei eventuellen Tubulysin Knockout-Mutanten wurden die Extrakte zunächst über einen HPLC-Lauf fraktioniert und anschließend die Fraktionen mittels Toxizitätstest auf Tubulysin untersucht. Durch die vorhergehende Fraktionierung wird eine Maskierung der Tubulysin-Wirkung durch Myxothiazol vermieden. Da die beiden

Sekundärmetabolite verschiedene Retentionszeiten bei der Elution von einer C-14 Säule besitzen, liegen beide im folgenden Toxizitätstest in unterschiedlichen Fraktionen vor.

1.3 Toxizitätstest

Die Mini-Kulturen aus den 96-Mikrotiterplatten wurden nach Kultivierung durch Stickstoff-Begasung auf einem Heizblock bei 37°C eingetrocknet. Danach wurden die Zellpellets über 2 h in 100µl Methanol resuspendiert und jeweils 10 µl für den folgenden Toxizitätstest eingesetzt, um eine Tubulysin-Produktion der jeweiligen Mutante detektieren zu können.

Für diesen Test werden L929 Zellen in DMEM-Medium (Invitrogen, Groningen) bei 37°C kultiviert und danach mittels Zellschaber vorsichtig geerntet. Diese Zellsuspension wird anschließend 1 : 10 mit DMEM verdünnt und 120 µl pro Öffnung einer 96-Mikrotiterplatte verteilt. Diese werden anschließend mit den 10 µl Zellextrakt der einzelnen Transposon-Mutanten versetzt und für fünf Tage bei 37°C inkubiert. Nach dieser Inkubationszeit werden die L929 Zellen mikroskopisch auf ZellkernfrAGMENTIERUNG untersucht, was ein Zeichen für Tubulysineinwirkung ist. Bei Zellen, die keine ZellkernfrAGMENTIERUNG zeigten, wurden die entsprechenden Mutanten als mutmaßliche Tubulysin Knockout-Mutanten identifiziert. Die Extrakte dieser Mutanten wurden in 50 ml M7-Medium (+ 1ml Absorber-Harz XAD-16 von Rohm & Haas) angezogen und die Zellkerne der L929-Zellen nach durchgeführtem Toxizitätstest zusätzlich durch eine Anfärbung des Chromosoms mittels DAPI-Färbung auf eine ZellkernfrAGMENTIERUNG bzw. Tubulysin-Produktion untersucht.

1.4 Bestimmung des Integrationsgenortes der Tubulysin Knockout-Mutanten in der An d48 mariner vermittelten Mutantenbank mittels Transposon-recovery

In der generierten Mutantenbank konnten mittels Toxizitätstest fünf Mutanten identifiziert werden (MutT176, 524, 781, 794 und 929), die kein Tubulysin produzierten. Dieses Ergebnis konnte nach Rekultivierung der Mutanten aus der Dauerkultur und erneuter Analyse bestätigt werden. Um Informationen zu erhalten, in welchen Bereich des Genoms das *Himar1* Element transponiert ist, wurde ein Transposon-recovery durchgeführt. Bei dieser Methodik wird die chromosomale DNA der jeweiligen Mutante mit verschiedenen Restriktionsenzymen geschnitten, die nicht innerhalb der bekannten *magellan4* Sequenz schneiden. Die restringierte DNA wird ligiert, nach Transformation in DH5 α / λ pir-Zellen wird auf Kanamycin haltigen LB-Platten bei 37°C inkubiert. Auf diesen Platten können nur solche *E.coli* Zellen wachsen, die ein Plasmid mit dem *magellan4* und somit dem Tn5 Kanamycin-Resistenzgen enthalten. Solch ein Plasmid enthält an den Enden des Transposons chromosomale DNA aus An d48. Diese Plasmide können in den *E.coli* Zellen DH5 α / λ pir replizieren, da innerhalb der Transposon Sequenz der *oriR6K* sitzt. Das Transposon wurde so aus dem jeweiligen Genom isoliert und mit den Primern K388 und K389 ansequenziert. Die erhaltenen Sequenzen wurden dann gegen die Genbank auf Homologien zu bekannten Genen hin untersucht und zeigten dabei hohe Ähnlichkeiten zu nicht ribosomalen Peptidsynthetasen (NRPS) aus bekannten Sekundärmetabolit-Biosynthesegenclustern wie denen des Myxothiazols, Nostopeptolids und Saframycins. Diese Analysen waren eindeutige Hinweise dafür, dass es sich dabei um Sequenzfragmente aus dem gesuchten Tubulysin-Gencluster handelt. Durch Restriktions- und Southernanalysen wurde die Größe der

einzelnen Transposon-Plasmide und ihre relativen Integrationsorte zueinander (innerhalb des Genclusters) bestimmt.

1.4.1. Transposon-recovery

Isolieren chromosomaler DNA nach Standardprotokollen aus 50 ml Tryptonmedium-Kultur jeder *A. disciformis* An d48 Mutante. Von dieser DNA werden 5 µg für die folgende Ausklonierung des Transposons verwendet, indem zunächst eine Restriktion durchgeführt wird. Dabei wurden die Enzyme *NotI* und *BamHI* verwendet, die keine Restriktionsstelle innerhalb des *magellan4* haben und statistisch relativ häufig in GC-reicher DNA schneiden sollten.

Verdau der genomischen DNA mit *NotI* bzw. *BamHI*:

5 µg DNA
+ 3 µl 10 x NEB Puffer
+ 3 µl 100 x BSA
+ 10 U Restriktionsenzym (*BamHI* bzw. *NotI*)
+ x µl dest. H₂O
30 µl Ansatz für 3 h bei 37°C inkubieren ⇒ wiederum 10 U Enzym zu den Restriktionsansätzen zugeben und weitere 2 h bei 37°C inkubieren.

Fällen der restringierten DNA und folgende Ligation

Der gesamte Restriktionsansatz von 30 µl wird mit 1 Vol. Chloroform/Phenol versetzt und für 10 min. zentrifugiert (13.200 rpm; 20°C). Der Überstand wird in ein neues Reaktionsgefäß überführt und mit 1/10 Vol. 3 M NaOAc und 2,5 Vol. 100% EtOH versetzt. Zum fällen der DNA wird das

Reaktionsgefäß für 1 h bei -20°C inkubiert und anschließend für 30 min. zentrifugiert (13.200 rpm; 4°C). Der Überstand wird verworfen und das Pellet dreimal mit 70% EtOH gewaschen, wobei man jeweils 5 min. zentrifugiert (13.200 rpm; 20°C). Nach dem Verwerfen des Überstandes wird das Pellet bei 37°C getrocknet und in 15 µl H₂O resuspendiert. Für die anschließende Ligation werden die gesamten 15 µl der gefällten DNA verwendet.

Ligationsansatz:

	15 µl	DNA	
+	4 µl	5 x Ligasepuffer (NEB)	
+	<u>1 µl</u>	<u>NEB Ligase</u>	
	20 µl	Ansatz über Nacht bei 16°C inkubiert	

⇒ wiederum 1 µl Ligase zu den Ligationsansätzen und ÜN
bei 16°C inkubieren.

Elektrotransformation der Ligationsansätze in den *E. coli* Stamm DH5 α - λ pir

1-3 μ l der Ligationsansätze und 50 μ l DH5 α - λ pir Zellen werden gemischt und in eine auf Eis gekühlte Elektroporationskuvette (0,1 cm) gegeben. Die Elektroporation wird bei 200 Ω , 25 mF und 1,25kV / cm durchgeführt. Die Zellen werden danach in 1 ml LB-Medium (10 g Trypton; 10 g NaCl; 5 g Hefeextrakt auf 1 l Medium) aufgenommen und für 1 h bei 37°C inkubiert. Danach werden sie auf Kanamycin-haltigen (50 μ g / ml) LB-Platten ausplattiert. Nach einem Tag Inkubation bei 37°C können die Klone gepickt werden. Dabei können nur Zellen anwachsen, die ein Transposon-Plasmid beinhalten und damit eine Tn5-Kan^R vermittelte Resistenz besitzen.

1.5 Sequenzauswertung des Tubulysin Biosynthese-Genclusters aus pMutT794/NotI

Das Transposon-Plasmid pMutT794/NotI enthält 52985 bp chromosomale DNA aus *Angiococcus disciformis* An d48. Zusammen mit dem *HimarI* Minitransposon *magellan4* (2199 bp), das bei Basenpaar 37317 bp in das Plasmid integriert ist, wurden 55184 bp sequenziert. Insgesamt 21760 bp stammen aus codierenden Genen des Tubulysin-Genclusters und sowie 31219 bp weitere codierende Gene. Bei diesen ORF's handelt es sich zum Teil um Regulatorgene, die die Expression des Tubulysins beeinflussen können. Sequenzvergleiche mit den Transposon-Plasmiden der anderen Tubulysin Knockout-Mutanten zeigten, dass das *magellan4* bei den Mutanten MutT781 (36975bp) und MutT929 (36197 bp) innerhalb von 1658 bp zu MutT794 in das Biosynthese-Gencluster transponiert sind.

Auf der Sequenz ist der Start des Tubulysin-Genclusters mit drei NRPS-Modulen (*tubA-C*), einem Cyclodeaminase codierendem Gen (*tubZ*) und einem PKS-Modul (*tubD*) enthalten. Desweiteren liegen innerhalb des Genclusters ein Anionentransporter codierendes Gen (ORF1), das zum Austransport des Tubulysins aus der Zelle dient und ein weiterer ORF (ORF2). Die grundsätzliche Anordnung der Gene, sowie der einzelnen Domänen mit einer N-Methyltransferase innerhalb der Adenylierungsdomänen (A) von *tubB* und *tubC*, entsprechen dem typischen Aufbau des Genclusters und der damit verbundenen Tubulysin-Biosynthese. Die Methyltransferase-Domänen (NMT) sind aber im Gegensatz zu bekannten Genclustern-Strukturen nicht zwischen der Adenylierungs- und Thiolierungs-Domäne (PCP) lokalisiert, sondern zwischen A8 und A9 innerhalb der Adenylierungs-Domäne (A) (hoch konservierte Regionen innerhalb der Adenylierungs-Domänen von NRPS; Konz & Marahiel, Chem. Biology, 6 (1999) R39 - R48). *TubA*, codiert eine unvollständige

Kondensationsdomäne, die für die Biosynthese theoretisch nicht benötigt wird. Die am Ende der bekannten Sequenz liegende Polyketidsynthase (PKS) enthält eine Ketoacylsynthase- (KS), Acyltransferase- (AT) und Ketoreduktase- (KR) Domäne.

Die restliche Sequenz des Tubulysin-Biosynthesegenclusters wurde aus einer Cosmidbank von An d48 (hinterlegt bei der DSMZ) unter Standardbedingungen identifiziert. Das auf der ersten Hälfte der Sequenz endende PKS-Modul (*tubD*) wird fortgesetzt von den schon erwähnten KS, AT und KR-Domänen und beinhaltet weiterhin eine Enoylreduktase (ER) und ein Acyl Carrier Protein (ACP). In der folgenden Sequenz von *tubD* wird eine NRPS codiert, die eine Heterozyklisierungs- (HC), Adenylierungs- (A) und Peptidyl Carrier Protein- (PCP) Domäne trägt. Des weiteren folgen die Gene *tubE* und *tubF*. Das Gen *tubE* codiert eine NRPS mit den Domänen C, A und PCP. Auf *tubF* wird eine PKS mit der folgenden Domänen-Anordnung codiert: Ketoacylsynthase (KS), Acyltransferase (AT), Ketoreduktase (KR), C-Methyltransferase (CMT), Dehydratase (DH), Enoylreduktase (ER), Acyl Carrier Protein (ACP) und schließlich eine Thioesterase, die für die Abspaltung des fertigen Tubulysins in Form einer freien Säure vom Multienzymkomplex sorgt. Der Insertionsort des Transposons *magellan4* liegt bei der MutT176 bei Basenpaar 54579 innerhalb des Biosynthesegenclusters. Der Insertionsort der Mutante MutT524 liegt nicht auf der uns bekannten Gencluster-Sequenz. Wir postulieren daher, dass der Insertionsort innerhalb eines für eine Acyltransferase codierenden Genes liegt, das downstream vom Tubulysin-Biosynthesegencluster liegt und eine posttranslationale Funktion zur Modifikation des Tubulysins hat.

2. Identifizierung der Anschlusssequenz des Tubulysin-Biosynthese-Genclusters aus *Angiococcus disciformis* An d48

2.1 Identifizierung und Charakterisierung von Cosmiden, die eine überlappende Sequenz downstream zum Tubulysin-Biosynthese-Gencluster tragen

Im vorhergehenden Beispiel wurde beschrieben wie die erste Hälfte des Tubulysin-Biosynthese-Genclusters zusammen mit weiteren an der Biosynthese beteiligten Genen mittels mariner basierender Transposon-Mutagenese und anschließendem Transposon-recovery identifiziert und annotiert wurde. Da innerhalb dieser Sequenz sowohl codierende Gene für Monooxygenasen als auch Acyltransferasen fehlen, sollte noch weitere Sequenz downstream von dieser identifiziert und charakterisiert werden. Innerhalb dieser Sequenz sollten oben angegebene Gene codiert vorliegen, da sie für die Biosynthese des Tubulysins erforderlich sind. Das Biosynthese-Gencluster sollte auf diesem Weg vollständig identifiziert werden.

Hierfür wurde eine Cosmidbank aus *A. disciformis* An d48 mittels Gigapack II XL packaging-kit (Fa. Stratagene) in *E. coli* SURE erstellt. Innerhalb dieser Bank sollten Cosmide identifiziert werden, die einen längeren Überlapp mit dem Tubulysin-Biosynthese-Gencluster downstream von *tubF* besitzen. Dafür wurden aus der bekannten Sequenz des Tubulysin-Genclusters zwei Primerpaare abgeleitet und die PCR-Amplifikate als Sonden für die folgende Hybridisierung der Cosmidbank eingesetzt. Das erste Primerpaar ASTls1A-1B liefert ein 889 bp großes DNA-Fragment und liegt 1 kb vor der *NotI*-Restriktionsschnittstelle in *tubD*. Das zweite Primerpaar ASTls2A-2B generiert ein 700 bp Fragment, das 11 kb upstream im bekannten Cluster-Ende entfernt in *tubC* liegt.

Die PCR wurde bei 54°C Annealing-Temperatur durchgeführt. Durch diese Hybridisierung konnten verschiedene Cosmide identifiziert werden. Mittels PCR- und Restriktionsanalytik wurden sie auf die Größe ihrer Überlappung mit der bekannten Cluster-Sequenz untersucht. Hierfür wurden wiederum die Primerpärchen ASTls1A/B und ASTls2A/B bei einer Annealing-Temperatur von 58°C eingesetzt. Bei der Restriktion wurden verschiedene Enzyme in Einzel- und Doppelrestriktionen eingesetzt.

Da die Cosmide F7 und F13 nach den Restriktionsanalysen einen ähnlich großen Überlapp mit dem ersten Teil des Clusters zeigten, trägt eines dieser Cosmide die nötigen genetischen Informationen, um die unmittelbar clustergebundenen Gene identifizieren zu können.

2.1.1 Southern-Analytik der Cosmide F7 und F13

Für die Identifizierung des richtigen Cosmids wurden zunächst Restriktionsenzyme ausgewählt, die möglichst selten und am Ende der bekannten Genclustersequenz schneiden. Die ausgewählten Enzyme waren *NdeI* und *NsiI*, die an den Positionen 39306 bp bzw. 39430 bp schneiden. Des weiteren schneiden beide Enzyme nur ein weiteres Mal in der bekannten Sequenz. Mit einer generierten Sonde (Primerpaar Tls_{up/down}), die hinter diesen Schnitten direkt am Ende der bekannten Cluster-Sequenz bindet, sollte nun die Cosmid-Genbank „gescreent“ werden. Die Cosmide wurden dafür in verschiedenen Doppelrestriktionsansätzen hydrolisiert und auf einem Agarosegel (0,8 %) aufgetrennt. Für die Doppelrestriktion wurden neben *NdeI* und *NsiI* die Enzyme *BamHI*, *EcoRI*, und *NotI* ausgewählt. Durch die Kombinationen mit *EcoRI* und *NotI* sollte erreicht werden, dass mittels der Hybridisierung ein Fragment identifiziert wird, welches bis an das Ende des jeweiligen

Cosmid-Inserts reicht. Falls dieses Fragment für eine nachfolgende Klonierung zu groß sein sollte, wurde ebenfalls *Bam*HI verwendet, um eventuell kürzere Fragmente zu erhalten. Die Hybridisierung wurde bei 42°C durchgeführt, und es wurde unter höchst stringenten Bedingungen gewaschen (68°C).

Das Ergebnis dieser Analytik war, dass in dem *Eco*RI / *Nde*I-Restriktionsansatz des Cosmids F7 ein 12 kb Fragment detektiert wurde. Dieses Fragment enthält die restliche Sequenz des Tubulysin-Genclusters und reicht bis zum Ende der Insert-Sequenz des Cosmids. Diese Schlussfolgerung wurde aus den Restriktionsanalysen und der charakterisierten Überlappung mit der Tubulysin-Genclustersequenz geschlossen. Das detektierte *Not*I / *Nde*I - Fragment ergab eine Größe von 4,2 kb. Daher muss innerhalb der 12 kb Insertsequenz von F7 - zwischen Anfang (*Nde*I - Restriktionsstelle) und Vektor (*scos*) - mindestens eine weitere *Not*I-Schnittstelle liegen. Somit kann die Anschlusssequenz in kleineren Fragmenten (als *Nde*I / *Not*I - und *Not*I / *Not*I-Fragmente) kloniert und sequenziert werden. Der *Bam*HI / *Nsi*I Doppelrestriktionsansatz lieferte insgesamt fünf Fragmente.

2.1.2. Klonierung der restlichen Tubulysin-Genclustersequenz aus Cosmid F7

Das Cosmid F7 wurde in einem Doppelrestriktionsansatz mit den Restriktionsendonucleasen *Nsi*I und *Eco*RI geschnitten (2 h; 37°C). Nach einer Auftrennung des Restriktionsansatzes mittels 0,8 %igem Agarosegel, wurde die entsprechende Bande aus dem Gel ausgeschnitten und mit dem NucleoSpin kit (Fa. Macherey-Nagel) extrahiert. Das isolierte Fragment wurde mit *Not*I

nachgeschnitten, um zu überprüfen, ob die erzielten Hybridisierungsergebnisse bestätigt werden. Zusätzlich wurde überprüft ob weitere *NotI*-Erkennungssequenzen innerhalb der 12 kb Anschlusssequenz liegen, um die Anzahl der zu klonierenden Teilfragmente bestimmen zu können. Die Restriktion ergab nach einer Auftrennung auf einem 0,8%igem Agarosegel ein 4,2 (*NsiI* / *NotI*-Fragment) und 8 kb großes Fragment (*NotI* / *NotI*-Fragment).

Zum einen konnte somit das Hybridisierungsergebnis bestätigt werden und beide Fragmente konnten für eine Klonierung eingesetzt werden. Zum anderen wurde bestätigt, dass es sich um das richtige Fragment handelt, das die Sequenz downstream des Clusters trägt.

Kloniert wurden das 12,2 kb große *NsiI* / *EcoRI* - Fragment, sowie die 4,2 kb (*NsiI* / *NotI*) und 8 kb (*NotI* / *NotI*) großen Fragmente in den Vektor pUC18. Der Vektor wurde für die folgende Ligation mit *PstI* / *EcoRI* bzw. *PstI* / *NotI* und *NotI* geschnitten. *PstI* und *NsiI* besitzen ein kompatibles Schnittmuster, so dass nach erfolgreicher Klonierung, diese Schnittstellen nicht mehr vorhanden sind. Mittels *HindIII* bzw. *NdeI* und *EcoRI* kann in einer Doppelrestriktion das 12 kb Insert wieder aus dem pUC18 - Derivat herausgeschnitten werden.

Die erhaltenen Klone wurden mittels dieser Restriktionsansätze auf ihre Richtigkeit hin überprüft. Einer dieser Klone (ASpUC12) wurde für eine folgende *in vitro* Transposition mittels GPSTM - 1 Genome Priming System (Fa. New England Biolabs_{Inc}) eingesetzt.

2.1.3. In vitro Transposition mit GPS - 1TM Genome Priming System

Durch das GPS - 1 System sollte mittels einer auf Tn7 basierenden *in vitro* Transposition das klonierte *NsiI* / *NotI* Fragment sequenziert werden. Bei diesem „kit“ wird eine TnsABC Transposase verwendet, die das Transposon (Transprimer™) zufällig in die Zielsequenz inseriert. Durch spezifische Sequenzierprimer (PrimerN / PrimerS), die aus den flankierenden Enden des Transposons „herauslesen“, können die anliegenden Bereiche des DNA-Inserts sequenziert werden. Da das Transposon zufällig in die Zielsequenz inseriert wird, kann durch Sequenzieren einer bestimmten Anzahl generierter Transposon-Mutanten die gesamte Zielsequenz charakterisiert werden.

Durchführung der *in vitro* Transposition

	2 µl 10 x GPS-Puffer
+	1 µl pGPS 1.1 (vermittelt Kan ^R)
+	0,2 µl Ziel - DNA (entsprechen 80 ng ASexp7)
+	14,8 µl dH ₂ O
	<hr/>
	18 µl Ansatz

Dieser Ansatz wird gut durchmischt und mit 1 µl TnsABC Transposase versetzt (wiederum durchmischen). Der gesamten Reaktionsansatz wird für 10 min. bei 37°C inkubiert, damit sich die Transposase vor der eigentlichen Reaktion im Reaktionsansatz vermischt. Nach der Zugabe von 1 µl „start solution“ wird der Reaktionsansatz für eine Stunde bei 37°C inkubiert. In dieser Zeit kommt es zum Strangtransfer des Transposons in die Ziel-DNA. Danach wird die Reaktion durch Inkubation für 10 min. bei 75 °C abgestoppt. Von diesem Ansatz wurden 2 µl in *E.coli* DH10B transformiert und auf Kanamycin-haltigem Medium ausplattiert. Insgesamt sind ca. 2000 Klone nach einer Inkubation über Nacht bei 37°C gewachsen.

Von diesen Klonen wurden 20 daraufhin überprüft, in welchem Verhältnis das Transposon in das Insert oder den Vektor inseriert ist. Dafür wurden diese Klone in einem Doppelrestriktionsansatz mit den Endonukleasen *EcoRI* und *HindIII* hydrolisiert. Die Restriktionsanalyse ergab, dass bei 75% der Klone das Transposon in das Insert inseriert waren. Somit wurden 192 Klone sequenziert, wodurch eine ca. 12 fache Abdeckung der Sequenz erreicht wurde (bei 500 bp Lauflänge je Sequenzierung).

2.2 Sequenzanalyse und Annotieren der 12 kb Anschlusssequenz

Die erhaltene Restsequenz des Tubulysin-Biosynthese-Genclusters ist 12.219 bp lang und hat eine Überlappung mit der vorher identifizierten Sequenz von 133 bp. Sequenzabschnitte, bei denen nur eine einfache Abdeckung vorlag, wurden durch erneute Sequenzierungen spezifischer Klone doppelsträngig sequenziert. In dieser Sequenz ist von Basenpaar 6416 - 6898 eine Acyltransferase (Position 76.787 - 77.545 bp in Gesamtsequenz) codiert. Die anderen identifizierten ORFs haben ebenfalls eine Funktion in der Tubulysin-Biosynthese. Die Gesamtsequenz beträgt somit 82.868 bp.

3. Identifizierung einer Tubulysin überproduzierenden Mutante innerhalb der mariner-Transposon Mutantenbank

Um die Mutanten der Transposonbank auf weitere auffällige Phänotypen gegenüber dem Wildtypen zu untersuchen, wurde eine HPLC-Analytik durchgeführt. Dabei wurde überprüft, ob die Insertion des Transposons in chromosomale Bereiche von

Biosynthese-Genclustern anderer, exprimierter Sekundärstoffe stattgefunden hat. Bei diesen Vergleichen gegenüber einem Extrakt des Wildtypen sollten dann nicht produzierende Mutanten der jeweiligen Metabolite identifiziert werden. Zu diesen Metaboliten gehören das Myxothiazol (Gerth et al. 1980 J Antibiot (Tokyo) 33(12):1474-1479 und Silakowski et al. 1999 J Biol Chem. 274(52):37391-9, Myxochelin (Gerth et al. 1983 J Antibiot (Tokyo) 36(9):1150-6. und Silakowski et al. 2000 Eur J Biochem. 267(21):6476-85) und Angiolam (Kunze et al. 1985, J Antibiot (Tokyo) 38(12):1649-54).

Bei den Auswertungen der 1.200 HPLC-Läufe fielen Extrakte einiger Mutanten auf, bei denen eine erhöhte Myxothiazol-Produktion gemessen werden konnte. Zur Kontrolle der erzielten Ergebnisse wurden 50 ml M7-Medium Kulturen der entsprechenden Mutanten erneut getestet und Zeitkinetiken zur Myxothiazol-Produktion über mehrere Tage gegenüber dem Wildtypen, erstellt. Die Ergebnisse dieser Untersuchungen zeigten eine deutlich erhöhte Produktion des Myxothiazols in den verschiedenen Mutanten bezogen auf den An d48 Wildtypen.

Die Bestimmung des Transposon-Insertionsortes innerhalb der jeweiligen Mutante erfolgte mittels „Transposon-recovery“ und anschließender Sequenzierung der flankierenden Bereiche (siehe 1.4). Die erhaltenen Sequenzen wurden gegen die Genbank auf Homologien zu bekannten Genen hin untersucht und zeigten hohe Ähnlichkeiten zu regulatorischen Elementen / Genen aus bakteriellen Organismen. Aufgrund dieser Ergebnisse wurde die gesamte Mutantenbank auf Tubulysin überproduzierende Mutanten hin untersucht. Dafür wurde der bestehende Toxizitätstest (siehe 1.3) optimiert. Durch mehrfache Verdünnungen des jeweilig eingesetzten Mutanten-Extraktes im Toxizitätstest wird eine

Ausdünnung des Tubulysins erreicht und somit ist die charakteristische Wirkung auf L929 Zellen ab einer bestimmten Verdünnung nicht mehr detektierbar. Anhand dieser Verdünnungsreihen (aus der gesamten Mutantenbank) wurden Mutanten identifiziert, bei denen signifikant höhere Verdünnungen benötigt werden, um keine Wirkung mehr feststellen zu können. Dies bedeutet, dass die jeweiligen Mutante eine erhöhte Produktion des Tubulysins aufweist.

Es konnte die Mutante Mut158 identifiziert werden, die eine vierfach erhöhte Tubulysin D-Produktion zeigte. Dieses Ergebnis wurde sowohl durch die Anzucht der Mutante in 50 ml Kulturen über HPLC-MS Untersuchungen, als auch mehreren Kinetiken mit folgendem, optimiertem Toxizitätstest gegenüber dem Wildtypen gezeigt. Mittels des Toxizitätstest wurde sogar eine achtfache Überproduktion von Tubulysinen ermittelt. Dabei wird die Gesamtwirkung aller Tubulysin-Derivate detektiert und nicht nur die von Tubulysin D. Die Mutante 158 weist somit völlig überraschend noch Überexpression von weiteren Tubulysinen gegenüber dem Wildtypen von *A. disciformis* auf. Dies war in keinster Weise zu erwarten gewesen. Die Ausklonierung des genomischen Bereichs unmittelbar an der Insertionsstelle des Transposons und die Sequenzierung erfolgte wie unter 1.4 beschrieben.

Die Sequenz des betroffenen Gens zeigt hohe Ähnlichkeiten zu einer Proteinkinase (aus *Stigmatella aurantiaca*), wobei der Insertionsort des Transposons die Promotorregion dieses Gens betrifft. Ohne auf diesen Wirkmechanismus festgelegt zu werden, besitzt dieses Gen eine negativ regulatorische Funktion für die Tubulysin-Bildung, weshalb eine Inaktivierung des Gens zu einer erhöhten Produktion führt. Die Gesamtsequenz beinhaltet 2.200

bp, wobei die Proteinkinase von Basenpaar 1.228 - 20 codiert wird und insgesamt 1.209 bp groß ist. Der upstream liegende ORF codiert ein Tubulysin-Biosynthese Protein und ist 933 bp groß.

Primersequenzen:

Sequenzierprimer für das *Himar1* Mini-Transposon *magellan4*

K-388: 5'→3' 5'TGG GAA TCA TTT GAA GGT TGG^{3'} SEQ ID NO: 39

K-389: 3'→5' 5'TGT GTT TTT CTT TGT TAG ACC G^{3'} SEQ ID NO: 40

Primerpaar ASTls1A/B wurde abgeleitet aus *tubD* und ergibt ein 889 bp Fragment

ASTls1A 5'CAC CCG GAC CTG CCT GGA TTC^{3'} SEQ ID NO: 41

ASTls1B 5'TGC TCG GCT GGC GCT ACT CAC^{3'} SEQ ID NO: 42

Primerpaar ASTls2A/B wurde abgeleitet aus *tubC* und ergibt ein 700 bp Fragment

ASTls2A 5'GCT CCC GGG CCA CGT GGT TGA AGA^{3'} SEQ ID NO: 43

ASTls2B 5'CCG CGG GCC GTG GCA GTG GTG TA^{3'} SEQ ID NO: 44

Primerpaar Tls_{up} / Tls_{down} wurde abgeleitet aus *tubF* und ergibt ein 125 bp Fragment

Tls_{up} 5'TGG CAG CCA GCC CGA GC^{3'} SEQ ID NO: 45

Tls_{down} 5'CCG CGG GTG CCC TCT CAT C^{3'} SEQ ID NO: 46

Name des Gens (ob PKS oder NRPS) Funktion SEQ ID NO der Aminosäuresequenz	Codierende Region in der Sequenz (SEQ ID NO:)	Codierende Region auf dem minus-Strang (SEQ ID NO:)	Größe in bp Größe des abgeleiteten/ übersetzten Proteins in Da	Klassifizierung der Domäne und Position in Nucleotidsequenz oder Protein mit höchster Ähnlichkeit
ORF16: Valyl tRNA-Synthase SEQ ID NO: 32	3.308 - 1 (SEQ ID NO: 1)	79.561 - 82.866 (SEQ ID NO: 18)	3.308 bp 122.434 Da	49% Identität, 64% Ähnlichkeit zu Valyl- tRNA Synthetase [<i>Thermotoga maritima</i>] NP_229614
ORF15: Regulatorische Komponente eines sensorischen Transduktionssystems SEQ ID NO: 31	4.706 - 3.453 (SEQ ID NO: 1)	78.163 - 79.416 (SEQ ID NO: 18)	1.254 bp 41.337 Da	38% Identität, 61% Ähnlichkeit zu "response" Regulator CheY Unterfamilie [<i>Synechocystis sp.</i>] NP_440346
ORF14: Zwei-Komponenten "response" Regulator SEQ ID NO: 2	5.719 - 7.164 (SEQ ID NO: 1)		1.446 bp 52.449 Da	29% Identität, 44% Ähnlichkeit zu "response" Regulator Protein [<i>Mesorhizobium loti</i>] NP_102571
ORF13: Zwei-Komponenten Regulationssystem aus Sensorkinase/"response" Regulator-Hybrid SEQ ID NO: 30	9.557 - 7.317 (SEQ ID NO: 1)	73.312 - 75.552 (SEQ ID NO: 18)	2.241 bp 83.814 Da	29% Identität, 46% Ähnlichkeit zu Zwei- Komponenten Regulationssystem aus Sensorkinase / "response" Regulator-Hybrid [<i>Agrobacterium tumefaciens</i>] NP_535879

ORF12: Hitzeschockprotein SEQ ID NO: 29	12.193 - 10.550 (SEQ ID NO: 1)	70.676 - 72.319 (SEQ ID NO: 18)	1.644 bp 58.650 Da	70% Identität, 85% Ähnlichkeit zu Hitzeschock - protein GroEL [<i>Rhodothermus marinus</i>] AAD37976	
ORF11: Tubulysin Biosynthese Protein SEQ ID NO: 3	12.841 - 13.881 (SEQ ID NO: 1)		1.041 bp 37.818 Da	34% Identität, 56% Ähnlichkeit zu hypothetischem Protein [<i>Corynebacterium glutamicum</i>] NP_616546	
ORF10: Tubulysin Biosynthese Protein SEQ ID NO: 28	14.833 - 13.835 (SEQ ID NO: 1)	68.036 - 69.034 (SEQ ID NO: 18)	999 bp 36.441 Da	33% Identität, 47% Ähnlichkeit zu konserviert hypothetischem Protein [<i>Streptomyces coelicolor</i>] NP_631315	
ORF9: Transkriptionsregulator SEQ ID NO: 4	14.942 - 15.586 (SEQ ID NO: 1)		645 bp 23.789 Da	30% Identität, 51% Ähnlichkeit zu Transkriptionsregulator, TetR Familie [<i>Caulobacter crescentus</i>] NP_420005	
ORF8: Integrase SEQ ID NO: 5	15.847 - 16.983 (SEQ ID NO: 1)		1.137 bp 41.614 Da	24% Identität, 40% Ähnlichkeit zu Integrase [<i>Corynebacterium glutamicum</i>] NP_601233	

ORF7:	21.154 - 18.809 (SEQ ID NO: 1)	61.715 - 64.016 (SEQ ID NO: 18)	2.346 bp 89.255 Da	33% Identität, 47% Ähnlichkeit zu hypothetischem Protein [<i>Nostoc sp.</i>] NP_490333	
Tubulysin Bioynthese Protein SEQ ID NO: 27					
ORF6:	22.366 - 23.532 (SEQ ID NO: 1)		1.167 bp 43.228 Da	32% Identität, 46% Ähnlichkeit zu Serin/Threonin-Kinase Pkn14 [<i>Myxococcus xanthus</i>] AAK64427	
Serin/Threonin-Kinase SEQ ID NO: 6					
ORF5:	24.591 - 26.513 (SEQ ID NO: 1)		1.923 bp 68.825 Da	42% Identität, 56% Ähnlichkeit zu Proteinkinase [<i>Stigmatella aurantiaca</i>] CAD19078	
Proteinkinase SEQ ID NO: 7					
ORF4:	26.597 - 27.517 (SEQ ID NO: 1)		921 bp 33.507 Da	33% Identität, 37% Ähnlichkeit zu Adenindeaminase- verwandtem Protein [<i>Deinococcus radiodurans</i>] NP_285591	
Adenindeaminase SEQ ID NO: 8					
ORF3:	29.858 - 30.400 (SEQ ID NO: 1)		543 bp 20.880 Da	24% Identität, 41% Ähnlichkeit zu möglicher Cytosindeaminase [<i>Salmonella typhimurium</i>] NP_462244	
Cytosindeaminase SEQ ID NO: 9					
<i>tubA</i> SEQ ID NO: 10	31.220 - 32.392 (SEQ ID NO: 1)		1.173 bp 43.202 Da	Ähnlichkeit zu C-Domänen "core motifs" / Kernmotive C2-C3	

ORF2	33.056 - 32.397 (SEQ ID NO: 1)	49.813 - 50.472 (SEQ ID NO: 18)	660 bp 25.390 Da	29% Identität, 42% Ähnlichkeit zu conserved hypothetical protein [<i>Neurospora crassa</i>] CAD11370	
SEQ ID NO: 26 <i>tubZ</i>	34.195 - 33.074 (SEQ ID NO: 1)	48.674 - 49.795 (SEQ ID NO: 18)	1.122 bp 40.499 Da	39% Identität, 53% Ähnlichkeit zu Lysin- Cyclodeaminase [<i>Streptomyces hygroscopicus</i>] CAA60467	
ORF1	35.422 - 34.205 (SEQ ID NO: 1)	47.447 - 48.664 (SEQ ID NO: 18)	1.218 bp 46.030 Da	40% Identität, 65% Ähnlichkeit zu Anionen transportierende ATPase [<i>Aquifex aeolicus</i>] NP_213468	
SEQ ID NO: 24 <i>tubB</i>	35.522 - 40.147 (SEQ ID NO: 1)		4.626 bp 170.704 Da	C 35.747 - 36.769 A 37.184 - 39.817 NMT 38.369 - 39.730 in A-Domäne zwischen "core motif" / Kernmotiv A8 und A9 PCP 39.818 - 40.069	76 - 416 555 - 1.432 950 - 1.403
SEQ ID NO: 11					1.433 - 1.516

<i>tubC</i> SEQ ID NO: 34	40.144 - 48.021 (SEQ ID NO: 1) 1 - 7.878 (SEQ ID NO: 33)		7.878 bp 289.141 Da	C 40.372-41.397 A 41.824 - 43.215 PCP 43.216 - 43.461 C 43.552 - 44.574 A 44.980 - 47.631 NMT 46.153 - 47.547 in A- Domäne zwischen "core motif" / Kernmotiv A8 und A9 PCP 47.632 - 47.868	77-418 561 - 1.024 1.025 - 1.106 1.137 - 1.477 1.613 - 2.496 2.004 - 2.468 2.497 - 2.575
<i>tubD</i> SEQ ID NO: 12	48.011 - 58.558 (SEQ ID NO: 1)		10.548 bp 383.778 Da	KS 48.011 - 49.321 AT 49.622 - 50.584 KR 51.473 - 52.309 ER 53.066 - 53.980 ACP 54.158 - 54.460 HC 54.461 - 55.870 A 56.000 - 57.412 PCP 57.413 - 57.643	1-437 538 - 858 1.155 - 1.433 1.686 - 1.990 2.050 - 2.150 2.151 - 2.620 2.664 - 3.134 3.135 - 3.211
<i>tubE</i> SEQ ID NO: 35	58.551 - 62.096 (SEQ ID NO: 1) 18.408 - 21.953 (SEQ ID NO: 33)		3.546 bp 130.337 Da	C 58.689 - 59.714 A 60.156 - 61.697 PCP 61.698 - 61.967	47 - 388 536 - 1.049 1.050 - 1.139
<i>tubF</i> SEQ ID NO: 13	62.103 - 70.616 (SEQ ID NO: 1)		8.514 bp 309.369 Da	KS 62.127 - 63.320 AT 63.711 - 64.676 KR 64.959 - 65.882 CMT 65.985 - 67.061 DH 67.242 - 67.829 ER 68.247 - 69.128 PCP 69.360 - 69.605 TE 69.759 - 70.586	9-406 537 - 858 953 - 1.260 1.295 - 1.653 1.714 - 1.909 2.049 - 2.342 2.420 - 2.501 2.553 - 2.828

ORF17: Ähnlichkeit zu Patatin ähnlichem Protein (Lipid Acylhydrolase) SEQ ID NO: 23	71.640 - 70.583 (SEQ ID NO: 1)	11.229 - 12.284 (SEQ ID NO: 18)	1.056 bp 38.371 Da	24% Identität, 40% Ähnlichkeit zu Patatin ähnlichem Protein [<i>Anabaena</i> <i>sp. 90</i>] CAC01602	
ORF18: Ähnlichkeit zu Patatin ähnlichem Protein (Lipid Acylhydrolase) SEQ ID NO: 22	72.786 - 71.731 (SEQ ID NO: 1)	10.083 - 11.138 (SEQ ID NO: 18)	1.056 bp 38.371 Da	24% Identität, 40% Ähnlichkeit zu Patatin ähnlichem Protein [<i>Anabaena</i> <i>sp. 90</i>] CAC01602	
ORF19: Tubulysin Bioynthese Protein SEQ ID NO: 21	75.209 - 74.655 (SEQ ID NO: 1)	7.660 - 8.214 (SEQ ID NO: 18)	555 bp 20.040 Da	31% Identität, 43% Ähnlichkeit zu hypothetischem Protein [<i>Azotobacter vinelandii</i>] ZP_00092207	
ORF20: Tubulysin Bioynthese Protein SEQ ID NO: 14	75.488 - 76.645 (SEQ ID NO: 1)		1.158 bp 43.282 Da	41% Identität, 62% Ähnlichkeit zu hypothetischem Protein [<i>Microbulbifer degradans</i>] ZP_00065421	
<i>tubG</i> : Acyltransferase SEQ ID NO: 15	76.787 - 77.545 (SEQ ID NO: 1)		759 bp 28.039 Da	47% Identität, 53% Ähnlichkeit zu N- hydroxyarylamine O- acetyltransferase [<i>Streptomyces avermitilis</i>] NP_826733	

ORF21: Tubulysin Bioynthese Protein SEQ ID NO: 16	77.769 - 78.695 (SEQ ID NO: 1)		927 bp 33.859 Da	28% Identität, 39% Ähnlichkeit zu konserviert hypothetischem Protein [Xanthomonas axonopodis] NP_641500	
ORF22: Tubulysin Bioynthese Protein SEQ ID NO: 17	79.138 - 80.019 (SEQ ID NO: 1)		882 bp 32.668 Da	37% Identität, 49% Ähnlichkeit zu hypothetischem Protein [Rhizobium etli] NP_659913	
ORF23: Tubulysin Bioynthese Protein SEQ ID NO: 20	81.319 - 80.057 (SEQ ID NO: 1)	1.550 - 2.812 (SEQ ID NO: 18)	1.263 bp 49.133 bp	34% Identität, 52% Ähnlichkeit zu hypothetischem Protein [Nostoc punctiforme] ZP_00109292	
ORF24: Carboxylatreductase SEQ ID NO: 19	82797 - 81.721 (SEQ ID NO: 1)	72 - 1.148 (SEQ ID NO: 18)	1.077 bp 37.621 Da	31% Identität, 44% Ähnlichkeit zu Pyrrolin Carboxylatreductase NosF [Nostoc sp.] AAF17284	
ORF25: Proteinkinase SEQ ID NO: 38	1.228 - 20	973 - 2181 (SEQ ID NO: 36)	1.209 bp 44.079 Da	Tubulysin überproduzierende Mutante	
ORF26: Tubulysin Bioynthese Protein SEQ ID NO: 37	2.157 - 1.225	44 - 976 (SEQ ID NO: 36)	933 bp 33.229 Da	Tubulysin überproduzierende Mutante	

KS: Ketoacylsynthase
AT: Acyltransferase
KR: Ketoreduktase
DH: Dehydratase
ER: Enoylreduktase
ACP: acyl carrier protein / Acyl-Trägerprotein
CMT: C-Methyltransferase
NMT: N-Methyltransferase
A: Adenylierungsdomäne
C: Kondensationsdomäne
PCP: peptidyl carrier protein / Peptidyl-Trägerprotein
TE: Thioesterase

bp: Basenpaare
Da: Dalton

Patentansprüche

1. ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:
 - (i) ss-DNA-Molekül mit einer Sequenz gemäß Figur 1;
 - (ii) ss-DNA-Molekül, das mit einem ss-DNA-Molekül gemäß (i) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zu jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von dem ss-DNA-Molekül gemäß (i) hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und

- (iii) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines ss-DNA-Moleküls gemäß (i) oder (ii) komplementär ist.
2. ds-DNA-Molekül aus einem ss-DNA-Molekül gemäß Anspruch 1 und einem dazu komplementären Strang.
3. ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:
- (i) ss-DNA-Molekül mit einer Sequenz der Positionen 3.308 bis 1 (ORF 16) der Sequenz gemäß Figur 1;
 - (ii) ss-DNA-Molekül mit einer Sequenz der Positionen 4706 bis 3453 (ORF 15) der Sequenz gemäß Figur 1;
 - (iii) ss-DNA-Molekül mit einer Sequenz der Positionen 5719 bis 7164 (ORF 14) der Sequenz gemäß Figur 1;
 - (iv) ss-DNA-Molekül mit einer Sequenz der Positionen 9557 bis 7317 (ORF 13) der Sequenz gemäß Figur 1;
 - (v) ss-DNA-Molekül mit einer Sequenz der Positionen 12193 bis 10550 (ORF 12) der Sequenz gemäß Figur 1;
 - (vi) ss-DNA-Molekül mit einer Sequenz der Positionen 12841 bis 13881 (ORF 11) der Sequenz gemäß Figur 1;
 - (vii) ss-DNA-Molekül mit einer Sequenz der Positionen 14833 bis 13835 (ORF 10) der Sequenz gemäß Figur 1;
 - (viii) ss-DNA-Molekül mit einer Sequenz der Positionen 14942 bis 15586 (ORF 9) der Sequenz gemäß Figur 1;

- (ix) ss-DNA-Molekül mit einer Sequenz der Positionen 15847 bis 16983 (ORF 8) der Sequenz gemäß Figur 1;
- (x) ss-DNA-Molekül mit einer Sequenz der Positionen 21154 bis 18809 (ORF 7) der Sequenz gemäß Figur 1;
- (xi) ss-DNA-Molekül mit einer Sequenz der Positionen 22366 bis 23532 (ORF 6) der Sequenz gemäß Figur 1;
- (xii) ss-DNA-Molekül mit einer Sequenz der Positionen 24591 bis 26513 (ORF 5) der Sequenz gemäß Figur 1;
- (xiii) ss-DNA-Molekül mit einer Sequenz der Positionen 26597 bis 27517 (ORF 4) der Sequenz gemäß Figur 1;
- (xiv) ss-DNA-Molekül mit einer Sequenz der Positionen 29858 bis 30400 (ORF 3) der Sequenz gemäß Figur 1;
- (xv) ss-DNA-Molekül mit einer Sequenz der Positionen 31220 bis 32392 (TubA) der Sequenz gemäß Figur 1;
- (xvi) ss-DNA-Molekül mit einer Sequenz der Positionen 33056 bis 32397 (ORF 2) der Sequenz gemäß Figur 1;
- (xvii) ss-DNA-Molekül mit einer Sequenz der Positionen 34195 bis 33074 (TubZ) der Sequenz gemäß Figur 1;
- (xviii) ss-DNA-Molekül mit einer Sequenz der Positionen 35422 bis 34205 (ORF 1) der Sequenz gemäß Figur 1;
- (xix) ss-DNA-Molekül mit einer Sequenz der Positionen 35522 bis 40147 (TubB) der Sequenz gemäß Figur 1;

- (xx) ss-DNA-Molekül mit einer Sequenz der Positionen 40144 bis 48021 (TubC) der Sequenz gemäß Figur 1;
- (xxi) ss-DNA-Molekül mit einer Sequenz der Positionen 48011 bis 58558 (TubD) der Sequenz gemäß Figur 1;
- (xxii) ss-DNA-Molekül mit einer Sequenz der Positionen 58551 bis 62096 (TubE) der Sequenz gemäß Figur 1;
- (xxiii) ss-DNA-Molekül mit einer Sequenz der Positionen 62103 bis 70616 (TubF) der Sequenz gemäß Figur 1;
- (xxiv) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii) oder (xxiii) unter stringenten Bedingungen hybridisierbar ist und insbesondere dieselbe Anzahl von Basen aufweist; und
- (xxv) ss-DNA-Molekül, das mit einem ss-DNA-Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii) oder (xxiii) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zu jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von diesem ss-DNA-Molekül hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und

(xxvi) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines Moleküls gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv) oder (xxv) komplementär ist.

4. ds-DNA-Molekül aus einem ss-DNA-Molekül gemäß Anspruch 3 und einem dazu komplementären Strang.

5. ss-DNA-Molekül, ausgewählt aus der folgenden Gruppe:

- (i) ss-DNA-Molekül mit einer Sequenz der Positionen 35747 bis 36769 (Domäne C des tubB-Gens) der Sequenz gemäß Figur 1;
- (ii) ss-DNA-Molekül mit einer Sequenz der Positionen 37184 bis 39817 (Domäne A des tubB-Gens) der Sequenz gemäß Figur 1;
- (iii) ss-DNA-Molekül mit einer Sequenz der Positionen 38369 bis 39730 (Domäne NMT des tubB-Gens) der Sequenz gemäß Figur 1;
- (iv) ss-DNA-Molekül mit einer Sequenz der Positionen 39818 bis 40069 (Domäne PCP des tubB-Gens) der Sequenz gemäß Figur 1;
- (v) ss-DNA-Molekül mit einer Sequenz der Positionen 40372 bis 41397 (Domäne C des tubC-Gens) der Sequenz gemäß Figur 1;

- (vi) ss-DNA-Molekül mit einer Sequenz der Positionen 41824 bis 43215 (Domäne A des tubC-Gens) der Sequenz gemäß Figur 1;
- (vii) ss-DNA-Molekül mit einer Sequenz der Positionen 43216 bis 43461 (Domäne PCP des tubC-Gens) der Sequenz gemäß Figur 1;
- (viii) ss-DNA-Molekül mit einer Sequenz der Positionen 43552 bis 44574 (Domäne C des tubC-Gens) der Sequenz gemäß Figur 1;
- (ix) ss-DNA-Molekül mit einer Sequenz der Positionen 44980 bis 47631 (Domäne A des tubC-Gens) der Sequenz gemäß Figur 1;
- (x) ss-DNA-Molekül mit einer Sequenz der Positionen 46153 bis 47547 (Domäne NMT des tubC-Gens) der Sequenz gemäß Figur 1;
- (xi) ss-DNA-Molekül mit einer Sequenz der Positionen 47632 bis 47868 (Domäne PCP des tubC-Gens) der Sequenz gemäß Figur 1;
- (xii) ss-DNA-Molekül mit einer Sequenz der Positionen 48011 bis 49321 (Domäne KS des tubD-Gens) der Sequenz gemäß Figur 1;
- (xiii) ss-DNA-Molekül mit einer Sequenz der Positionen 49622 bis 50584 (Domäne AT des tubD-Gens) der Sequenz gemäß Figur 1;

- (xiv) ss-DNA-Molekül mit einer Sequenz der Positionen 51473 bis 52309 (Domäne KR des tubD-Gens) der Sequenz gemäß Figur 1;
- (xv) ss-DNA-Molekül mit einer Sequenz der Positionen 53066 bis 53980 (Domäne ER des tubD-Gens) der Sequenz gemäß Figur 1;
- (xvi) ss-DNA-Molekül mit einer Sequenz der Positionen 54158 bis 54460 (Domäne ACP des tubD-Gens) der Sequenz gemäß Figur 1;
- (xvii) ss-DNA-Molekül mit einer Sequenz der Positionen 54461 bis 55870 (Domäne HC des tubD-Gens) der Sequenz gemäß Figur 1;
- (xviii) ss-DNA-Molekül mit einer Sequenz der Positionen 56000 bis 57412 (Domäne A des tubD-Gens) der Sequenz gemäß Figur 1;
- (xix) ss-DNA-Molekül mit einer Sequenz der Positionen 57413 bis 57643 (Domäne PCP des tubD-Gens) der Sequenz gemäß Figur 1;
- (xx) ss-DNA-Molekül mit einer Sequenz der Positionen 58689 bis 59714 (Domäne C des tubE-Gens) der Sequenz gemäß Figur 1;
- (xxi) ss-DNA-Molekül mit einer Sequenz der Positionen 60156 bis 61697 (Domäne A des tubE-Gens) der Sequenz gemäß Figur 1;

- (xxii) ss-DNA-Molekül mit einer Sequenz der Positionen 61698 bis 61967 (Domäne PCP des tubE-Gens) der Sequenz gemäß Figur 1;
- (xxiii) ss-DNA-Molekül mit einer Sequenz der Positionen 62127 bis 63320 (Domäne KS des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxiv) ss-DNA-Molekül mit einer Sequenz der Positionen 63711 bis 64676 (Domäne AT des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxv) ss-DNA-Molekül mit einer Sequenz der Positionen 64959 bis 65882 (Domäne KR des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxvi) ss-DNA-Molekül mit einer Sequenz der Positionen 65985 bis 67061 (Domäne CMT des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxvii) ss-DNA-Molekül mit einer Sequenz der Positionen 67242 bis 67829 (Domäne DH des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxviii) ss-DNA-Molekül mit einer Sequenz der Positionen 68247 bis 69128 (Domäne ER des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxix) ss-DNA-Molekül mit einer Sequenz der Positionen 69360 bis 69605 (Domäne PCP des tubF-Gens) der Sequenz gemäß Figur 1;

- (xxx) ss-DNA-Molekül mit einer Sequenz der Positionen 69759 bis 70586 (Domäne TE des tubF-Gens) der Sequenz gemäß Figur 1;
- (xxxi) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix) oder (xxx) unter stringenten Bedingungen hybridisierbar ist und insbesondere dieselbe Anzahl von Basen aufweist;
- (xxxii) ss-DNA-Molekül, das mit einem Molekül gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix) oder (xxx) hinsichtlich seiner Nucleotid-Anzahl oder seiner Nucleotid-Sequenz zur jeweils 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 oder 100% homolog ist, jedoch von diesem ss-DNA-Molekül hinsichtlich seiner Nucleotid-Anzahl und/oder seiner Nucleotid-Sequenz in mindestens einem Nucleotid abweicht; und
- (xxxiii) ss-DNA-Molekül mit einer Sequenz, die zur Sequenz eines Moleküls gemäß (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii), (xiv), (xv), (xvi), (xvii), (xviii), (xix), (xx), (xxi), (xxii), (xxiii), (xxiv), (xxv), (xxvi), (xxvii), (xxviii), (xxix), (xxx), (xxxi) oder (xxxii) komplementär ist.

6. ds-DNA Molekül aus einem ss-DNA-Molekül gemäß Anspruch 5 und einem dazu komplementärem Strang.
7. Varianten oder Mutanten, die aus einer Substitution, Insertion oder Deletion von Nucleotiden oder einer Inversion von Nucleotid-Segmenten eines ss-DNA-Moleküls gemäß Anspruch 1, 3 oder 5 oder eines ds-DNA-Moleküls gemäß Anspruch 2, 4 oder 6 resultieren, wobei diese Varianten und Mutanten Enzym-Varianten oder Enzym-Mutanten für die Produktion von Sekundarstoff(en) mit für Tubulysine charakteristischen Eigenschaften codieren.
8. RNA
 - (a) mit einer Sequenz entsprechend der eines ss-DNA-Moleküls gemäß einem der Ansprüche 1, 3, 5 oder 7 oder
 - (b) mit einer Sequenz einer RNA gemäß (a), aber in Gegenrichtung (anti-sense), oder
 - (c) mit einer Sequenz einer RNA gemäß (a) oder (b) und mit einem dazu komplementären Strang,jeweils gegebenenfalls als Element eines rekombinanten Vektors.
9. Rekombinanter Vektor, insbesondere Expressionsvektor mit einem DNA-Molekül gemäß einem der Ansprüche 1 bis 7.
10. Zelle, insbesondere zur Expression, in die ein DNA-Molekül gemäß einem der vorhergehenden Ansprüche oder einen Vektor gemäß Anspruch 8 oder 9 integriert ist.
11. Zelle nach Anspruch 10, wobei sich die Zelle von kultivierbaren Bakterien, insbesondere Myxobakterien, vorzugsweise Angiococcus, insbesondere A. disciformis, Archangium, insbesondere A. gephyra, Escherichia coli, Pseudomonaden oder Actinomyceten herleitet.

12. Verwendung eines Vektors gemäß Anspruch 8 oder 9 für die Transformation von Zellen oder Organismen zur transienten oder permanenten Expression eines oder mehrerer Proteine (Expressionsprodukt(e)), das (die) durch eine DNA (ssDNA oder dsDNA) des Vektors codiert wird (werden).
13. Verwendung einer Zelle gemäß Anspruch 10 oder 11 zur enzymatischen Biosynthese, Mutasyntese oder Partial-Synthese eines Tubulysins, insbesondere Tubulysin A, B; C, D, E und/oder F.
14. Expressionsprodukt eines DNA-Moleküls gemäß einem der Ansprüche 1 bis 7 oder eines Vektors gemäß Anspruch 8 oder 9 oder einer Zelle gemäß Anspruch 10 oder 11.
15. Polynucleotid enthaltend eine Sequenz wie in SEQ ID NO: 1, 18, 33 oder 36 definiert, oder ein Fragment davon.
16. Polynucleotid nach Anspruch 15, wobei das Fragment eine im Sequenzprotokoll definierte CDS ist.
17. Vektor enthaltend ein Polynucleotid nach Anspruch 15 oder 16.
18. Zelle enthaltend einen Vektor nach Anspruch 17.
19. Polypeptid enthaltend wenigstens eine Sequenz wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder ein Fragment und/oder Derivat davon.

20. Verwendung wenigstens eines Polynucleotids wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens eines Fragments davon und/oder wenigstens eines Polypeptids wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens eines Fragments davon zur Herstellung einer pharmazeutischen Zusammensetzung zur Behandlung von unerwünschtem Zellwachstums oder unerwünschter Zellvermehrung in einem Individuum.

21. Verwendung nach Anspruch 20, wobei das unerwünschte Zellwachstum oder die unerwünschte Zellvermehrung ein Tumor ist.

22. Verwendung nach Anspruch 20, wobei das unerwünschte Zellwachstum eine pathogene Infektion ist.

23. Verwendung nach Anspruch 22, wobei die pathogene Infektion eine Mykose, Malaria oder eine parasitäre Erkrankung ist.

24. Pharmazeutische Zusammensetzung, enthaltend wenigstens ein Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens ein Fragment davon und/oder wenigstens ein Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens ein Fragment davon.

25. Pharmazeutische Zusammensetzung nach Anspruch 24, die weiterhin wenigstens einen pharmazeutisch verträglichen Trägerstoff enthält.

26. Verfahren zur Herstellung von Tubulysinen und Tubulysin-Biosynthese Proteinen, umfassend die Schritte:

- (a) exprimieren von wenigstens einem Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder

wenigstens einem Fragment davon und/oder wenigstens einem Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens einem Fragment davon, und

(b) aufreinigen der Expressionsprodukte.

27. Verfahren nach Anspruch 26, wobei die Expression in prokaryotischen oder eukaryotischen Zellen und/oder durch in vitro Expression erfolgt.

28. Verfahren zum Auffinden von Genen, die an der Biosynthese von Tubulysinen beteiligt sind, umfassend die Schritte:

(a) hybridisieren von wenigstens einem Polynucleotid wie in SEQ ID NO: 1, 18, 33 und/oder 36 definiert und/oder wenigstens einem Fragment davon mit DNA, RNA, und/oder cDNA einer Spezies, die nicht mit *Angiococcus disciformis* identisch ist, und

(b) isolieren und charakterisieren der hybridisierten DNA, RNA, und/oder cDNA.

29. Kit zur Herstellung von Tubulysinen, enthaltend:

(a) wenigstens ein Polynucleotid nach Anspruch 15 oder 16 und/oder wenigstens einen Vektor nach Anspruch 17 oder

(b) geeignete Medien und Puffer zur Vermehrung von Zellen, die eine Expression des Polynucleotids und/oder Vektors erlauben und

(c) geeignete Mittel zum Aufreinigen des/der Expressionsprodukte(s).

30. Verwendung einer Zusammensetzung enthaltend wenigstens ein Polypeptid wie in SEQ ID NO: 2 bis 17, 19 bis 32, 34, 35, 37 und/oder 38 definiert und/oder wenigstens ein biologisch aktives Fragment oder Derivat davon als Desinfektionsmittel.

31. Verwendung nach Anspruch 30, wobei die Zusammensetzung flüssig oder pulverförmig ist.

32. Desinfektionsmittel wie in Anspruch 30 oder 31 definiert.

Figur 1

SEQUENCE LISTING

<110> Gesellschaft fuer Biotechnologische Forschung mbH

<120> Tubulysin-Biosynthesegene

<130> 13289

<160> 1

<170> PatentIn version 3.1

<210> 1

<211> 70782

<212> DNA

<213> *Angiococcus disciformis*

<400> 1

```
gcgggcgcct tcttcaccgc gggggcctcg gcctcctcgg tggtcggggc gggcgccagc      60
ttgtcgccgt gcaccttgcc gggagcggtc gccttcttct tcgaggcgga gcgcacgggc      120
tcgccttgcg tcacctcgcg cgtgggcagc atgtcccctg gggccttcac cgggcccggac      180
tcacctcctc ccgcctcggc gccctcggcc ggcgtctcgc gggtgtcgcc accctccttg      240
gcccgggtga actcgcgcat cgcgtcgtcc accagcccca tggtcagtga cgccacgccg      300
aggtcattgg ggtcggatgg cgacaggccc tccttcgtgc cctcgcgag cttcgcgagc      360
gcgtcctgca cctgcgtgtc ccgggcctcg gacgcgggtg cgggtctcgcc cagttcacc      420
ttcagctcct ggcccaggtc cagccgccc ttctcttctc gattcggggc gaccttcacc      480
tgggcccggg cggaggggga ggcacctct tccgcgggtg gggtgctgct ctcgggagtg      540
gtgctctcgg acggcgcggt ctctggcatg gggggctccg gggcaatccg ctgcagggtg      600
tcctgaagct tggccttgcg ctcttcagc tcctcgacgc gggcgcggtc cttctccacc      660
acgtccgggg gcgccttgcc caggaagttg ggggttctca gcttgcgag cagccggcc      720
atctcctgct cggcgcgggc aatctccttc ttcaaccgct cgcgctccgc gtccaggctg      780
acgaggcccc ccagcgggac gtagatctcc aggttggagc ccacgaaggc ggcggcctgc      840
ggcggtttgg cggccggcgc tcccaccgtc acctcggaca ggcgggccag cggcatgagg      900
tagccgcgcc agcgctccag cagctcgcgc gtgcgcacgt ccgcgctctg caccaccgcc      960
ttcaccttgg tggcgggcca caggttgctc tcgcgcgga tggtgcgag gccctcgatg     1020
gcggcgatga ccggcgccat ctgccttcc gccgcctcgt ccacgtgcgc cgcgtccggc     1080
tccgggtacg gggcaatcat gatgctgtcc gtcggccggg ccacggcag cttctgccag     1140
atctcctcgg tgatgaacgg catgaacggg tgcagcagcc gcaggatgcg gtccaggcag     1200
tacacgagca cccggcgcggt ggagtccttc gccgcgcgt cctcgccgta cagcgagccc     1260
ttcgccagct cgatgtacca gtcgcagaaac tccgcccaga ggaactggta gagcgtggag     1320
gcggcctcgc cgaagctgta cgactccagc agcgcgcgcg cctccgtggg ggcgcgctgg     1380
```

2/38

aggcgggaga gaatccagcg gtccgccagc gtcagtgtgc tccggtccag cgggccggcc	1440
tccatccgga agtcgcccac gtccatcagg gcgaagcggc tggcggtcca cagcttggtg	1500
ccgaaggcct tgtagccctc gagccggctc atcgacagct tgatgtcgcg gccctgctgg	1560
gtgagcgagg cgagcgtgaa ggcagcgcg tccgcgccga acggcggcac gccctgcggg	1620
aaccggttct tcagcgtcgg ctggagcttg tcggccgggg cgccgaggac gatgtccagg	1680
gggtcgatga cgttcccctt cgtcttggaac atcttctggc ccttctcgtc gcgcaccatc	1740
gcgtgcaggt acacgggtcg gaagggcaca tcccccatga agtggatgcc catcatcatc	1800
atccggggca ccagaagaa gatgatgtcg tggcccgctc ccatgacgga cgtcgggtag	1860
aaggtcttca gctccggcgt ctacgcggc cagcccagcg tggagaacgg ccacaggccg	1920
gacgagaacc aggtgtccag cacgtccggg tcctggatga aggacgcgcc gccgcacttc	1980
gggcaggact tcggctgctc gcgcgcgacg atgggctccg cgcgcgcgaa gtccaccccg	2040
cccaccttca ccgtgggggc gtccagcggc aggtccgtgt cgtcgccctg ccgcgggctg	2100
cacgaggtgc agtagtacgc gggaaatctg tggccccacc acagctggcg gctgacgcac	2160
cagtcgtgga tgttgcgcat ccagtggaag aacgtgttcg tccaggactc ggggacgaac	2220
ttcgtgcggc cctgctccac cgcctcgatg gccggcttcg ccagcggttc aatcttgatg	2280
aaccactgcg gagacaggcg tggctccacc accgtggcgc agcgtggca ggtgccacg	2340
ttcagcatgt ggggctcttc cttctccagc agccccctg cctgagggtc cgccagcacc	2400
tgcttgcgcg cctcgaagcg atccatgccg gcgtacttgc cggctctcct cgtcatccgg	2460
gccgcttcgt ccaggatggt gagcatgggc agcttggtgc gcaggcccg ctggtagtcg	2520
ttgaagtctg gcgcggcgt caccttcacc acgcgggtgc cgaacttcgg gtccaccagc	2580
tccgcgtccg cgatgatggg aatctcgcg tgggtcagcg gcagcttcac gctcttgccg	2640
gccagcccca ggtagcgctc gtcctccggg tggatggcca ccgcgggtgc gccagcatc	2700
gtctccgggc gtgtggtggc cacggtgagc gtgcggtcgc tgtccttgac ggggtagcgg	2760
atgtgccaga tggagccctt cttcgactcc tcgtgctcca cctccaagtc gctgaggggc	2820
gtgcggcacg aggggcacca gttgatgagc ttctggggcc ggtacatcag gccctcttcg	2880
tacaggcgca cgaagacctc gcgcacggcg gcggaggact gctcgtccat ggtgaagcgc	2940
tcgcggctcc agtccagcga ggcgccagg aagcggtgct gctcgccgat gcgggcgccg	3000
tacttgccct tccactccca gacgcgtcc aggaaggcgg cgcgcccgag gtcgtgccgg	3060
ctcttgccct cggctcttct cagctccttc tccaccacca tctgcgtggc gatgccggcg	3120
tgggtccgtgc cggggagcca cagggtgttg aagccgctca tccgcttcca gcgcgcgagg	3180
atgtcctgga tgggtggcgg gagcgcgtgg ccgatgtgga ggctgcccgt cacgttgggc	3240

ggcggcagca cgatggagaa ggcgggcttg tcggaggtcg cctcggcgcg gaagtagttc 3300
cgctccatcc agcaggcgta ccagcgggcc tcgacctcg tgggctcgta ggccttggac 3360
agttcagtgg tgctgggtcat tgcggacggc cagccccggg ggaaccgggg ccggcgggaa 3420
ggtcttggag agagggttaag gcgaggaacg cgtcagtgtc gcgtctcccg gtccttgatg 3480
agccgctcga gctcctcacg gatgatggtc tccgccagct gcggtaccac ctcccaggca 3540
atcttctcga tgacctcgcg ggaggccttc gagagcgcct cgcgcagcag cgcctcgcca 3600
ccatccgcgg cggtcggacg ggccgtggcc ggagcggcgg tcggcgccac cgcggtcgga 3660
gcggccgtcg gggggccgcg catgtccagg gagatctcct cggcaccgct gggctcgggc 3720
agcgagtcct cgatgctgat ggaaggctgg gctgcctgcg gcgcgcgct gggagcgccc 3780
aggccgaacg ggtccctcgc gcggggccgc ggctgggtcg cgccaccg cgcgggaggc 3840
agcggcgcg cgccggccgg aggacgcggg aagcccggt ggctgccctg gggaatcccc 3900
ggacggggcca tgccggcgcc ggcccacccg gaggcgcggg cagcccgga cgcgggcgca 3960
tgcccggcgc ggacgcaccc ggcggaggcg gcacgcccg gacgcgcggc ggcgcccgga 4020
ggcggcatgc ccggggcccg aggacgcggc gcgcccggag caccggcg cggcattcca 4080
ggccccggcg ggcgggcat gcccggtccc ggcggcattc caggcccccg aggacgcggc 4140
gcgcccggag caccggagg cggcattcca ggcccggag gacgggcat gccggcccc 4200
ggcggcattc caggcccccg aggacgcggc gcgcccggag caccggcg catgcccggg 4260
cccggaggac gcggcacgcc cggctgggtg ccctgcggaa tgcccgacg ggcggccgtg 4320
ggcgcgccc gctgcgggcc cggaggcacc ggcgcgggag tggccggtg ggcggtgggc 4380
gggcgcacct gcgtggccgc ggaggccggc atggtgttg acttctgccc caccagcgcc 4440
ttcaccttgt cgagcagcac ctggctctcg aagggttg cgatgtggtc atcggcgcgg 4500
gcggcgcggg cgcggttctc gtcgaaggcc tcgaaggctc ccgccagcag caccacgggg 4560
atgccctggg tggccgggtc gctcttcagg gcctcgcaga cctcgtagcc gctcttgccc 4620
ggcatcatca catcggcgag gacgacgtcc ggcgcagct cgcgggtacg ggagatggcg 4680
tccagcccgt tgtccaccgc ggtcacctga aagtcctcg tcgcgaagat catccaatc 4740
accttgcgga tggtagcgca gtcacggcg accagcagat tcttgggcat cggttcgggc 4800
ctcggggggc gtcccacccc cctgaattcg ttggagattc ctgccgcgaa aaaccaggt 4860
cgcgccctg tatcaagcg ggcagcttac gtttcgcgct ccagcacgat caagaaaaca 4920
tgcgctgcag gtccaggaac agcacggagc ctgtcaacc gggggcacgg aaggctccaa 4980
tggtgggtctc cggctcgaag tgctgcagca cgccgagcac ccgggtggcc gtcagcccca 5040
cgttccggcc ggccagctcc gtgaggatga agagggcctc ggggtaccgg ggggagccca 5100
gcagcgcggg gaccgagcag atggggccaca gcacctgggc gtgcgggaag atgccggcca 5160

ccggcccgt ctgcacgggc aggacgctga aggcctcccc tcgcgatatc acctgggaga 5220
cgaagcccag cgggacgccg aacaggcgcc cctgcgattc gaagacgagc gccttgctcg 5280
gcgtggactc ggcccagtgg acggggcgca gggccgcggg cgcggcacga gggcccaccc 5340
ggtggggcag cgcctcggcg atgagctcca ggtagagccg ctcttgctgc agcaccgcgc 5400
cccggctcag gggggccagc gagtccgcca gcccgggggg caggaggaag aacggcgccc 5460
gcgccacgtc ggcgacctcc accacggagc gcaccgcac ggccaggggtg ggactcacgt 5520
cgagcaccac caccatgccc ggccccctct cgggcaagcc gccgagcagg gccgacaggt 5580
ccttcacctc cagcacgcca cgcaggctgg agccgttggc gccgggcatg gccacttcca 5640
tgacggaggt ggcctccacc gcgtagcgcg tctctcctgc ctccacgaga aggcagagcc 5700
ggcgtccgct ttcgaagggt gacacggccc taccctagca gccttctctg ccttggtggt 5760
aagccatcga gccagatggc gcgcatcctc ctcgtcgacg acgaaaagat cgcccgacc 5820
ctctacggcg actacctcac ggccgtgggg cacgccgtca cggcgggtggg ctcgctccag 5880
gacgcgcggg aggcgctggc cggggaccgc ttcgacgcgg tggtagcgga cctcatcctc 5940
cccggcgggc acggcatgga ggtgctgcgc cacgtgcgcg agcgccaccc cggcgtggag 6000
gtggtggtca tcaccggcct ggagaagggtg gaccccgccg tgcgcgcat caagagcggc 6060
gcggccgagt acctcgtcaa gccggtggcc cccgaggccc tccagcatgc cgtccgccgt 6120
gccctacca cgcgagacct cctccaggag aacgcctcgc tgcgcggca cgtcgccatg 6180
ctggaggccg gccagcgcat cgccaccacg ctggagcgcg agaagctggc ctccgccacc 6240
acgagcgcgc tggagtccat ggcctgcgcc agcgccgtgg tgcgtctgga gcgcgacggc 6300
gacaagggcc tgcggctcca gggcgtccgc ggactgcctc cgcggcacga gccggagctg 6360
gtggccaccc tcgtcgagcg cctctcgaac gcgcgcgcgc cccgggagct ggaagggtg 6420
gacgccgct tcccgcgct cctctccatc cccgccgtgg agggcaccac ggtgctgggc 6480
cacgcggtgc tcttcttcga cggcgccgag ccggagtgga ccggcgagac ggcgggctac 6540
ctggtgcgca actgggcact ggcgctgcgc aacctcgggc gcttcgcggc ggtggaggac 6600
ctggcgtagc tcgacgacct caccgcctc ttcaacacgc gctacctgca cctggtgctg 6660
gaccgcgagg tccaggactc cgtccagacg cagcgccctc tcagcctgct gttcctggac 6720
ctggaccact tcaagtccat caacgacacg cacggccacc tcgtgggctc gcggctggtg 6780
gtggagacgg cccgggtggt gaagggtgc gtgagagacc acgacgtggt ggcccgtac 6840
ggcggcgacg agtacgtggt gctcctgcgc aacaccgact ccggcggcgc gctcaagggtg 6900
gcggagcgca tccgccgac catggagacg cacctgttcc tcgcgcgcga ggggctgtcg 6960
ctgaagctca gcacgtgcat cggcgtggcc agcttccccg agcacgcccg ggacaaggcc 7020

acgctgctgg acctggcgga cggggccatg taccggggca agcgtggctc ccggaatgtc 7080
gtctacatgg cggccctgga tctggagccg ctgccggcgg agcggcgcca gggcggcagc 7140
ggccccggcg gctccagcac gtagcggctc agcggcgagc gttgcccacg gtgaggcgct 7200
tggtcggggg ggtctccacc tcggcgcggtg ctcccccgcc ctcccttgccc accgcctcca 7260
cctcctcgtg cggggcaccg gtggtgatga gccgggtactt ctccgaggcc tcgcgggtcat 7320
gctgggcctg ctcggggtcc aaccgggcaa tcagctgcca gaagagggcg gcgtgctcgc 7380
gtcctcgtc catgacgtga cggaggatgg ccttggcctc ctcgttgtcg gtggcatcga 7440
tgtgggcggc gtagaggttg atggcgcca gctccgcctc gatgttcagc cggatggagc 7500
gggccagctc ggagtcggtc atcttccggg gcaccagcga gtggaacggg ttggtctgcg 7560
gcatgaaggc cctcccgagg ccggtgggtg cgtccacgtt ctggacgggg cgccagggga 7620
cgcgaggagc atccgcgcgg ccagggggcc ggggtcaacgg aatccctccg cccgcgttgc 7680
gtagataccg ctgcctccgc catgctcggg ccatggaaac accgcgcgcg ctctcccagt 7740
tgctccaggc cctggaggcg ggagatctgc cggccgcgcg cgcggcgggc gcggctctgc 7800
aacgtgcggg cgcgcactcg gcacagctgg cggcggagggt gctgcacgag ctgcggcagc 7860
ccctgctggg cgtgaaggcc tacgcgcagc tgctcgcgga ggacgggtggc cccaccggcc 7920
cgctgcggct gctgctggcg cagggtggagc ggatggaaca gatcgtctcc gactacatcc 7980
gcctcgccag cgagcggcac gcgccgcagc agcgcctgtc gctggcgggc cccatctggg 8040
cggcgggcga gctgttcagc gtcaaccggg actcggcgcg catgtctctg gaggtggagg 8100
cgcccgagga catcaccatc cagggcaacg cgcggctcat cgagcagctc acgctgaacc 8160
tgctgaacaa cgcgcgcgac gccatggccg gccgcggggc cgtgaagggtg gtgctctcgc 8220
gcgagggcgc ctctccgggtg ctgtacgtgg cggactgggg cccgggcatc cccgtggagc 8280
tgcgcgagcg catcttcgag ccctacgtca ccgccaacaa gcgcggcacg ggactggggc 8340
tgggcggtgt caagcgcctc gcccaggagc acaacgcgac gattggcctc gcggcgccgg 8400
gcgccattcg cgacgtgccc ccgccggcca ccgtgttccg ggtactcttc cccgccacgg 8460
acgcgcgcgc gcagctgcgc aagcgggtgc tgggtggtgga cgacgagacc atcatccgca 8520
tggtcttccg cgacctgatg ggcaaggagt gcgaggatcat cgaggcgggc agcggcgagg 8580
aagcgctgga tttgctgcgc caggcgccag tggacctcat cgtcacggac aagaacctgc 8640
cgggcctgtc cgggctggag ctggcgagc aggcgcggcg gctctactcc aactcgcgcg 8700
tcatectgat gacgggctac ccgtcgtgg tgacgacgca gcaggcgctg gagctgggag 8760
cctgctcaag cccttcgacg acatccggca ggtgcgcggg ctgttgcgca 8820
ctcggagccg ccggtgcctc cgggtggtgg ccccggggcc gaagtgcggc 8880
ggtggagcgt gctggaggac aaccgcagca cggcgcggt catctccgag gccctggaca 8940

6/38

tgctggggct ggaggcgcg gtgctgccct cgacggagct gatggcgatg gcgacgccgg 9000
tgggcgtggt ggtgagctgg gacttcgcgc cggcctatgg gcgcaaggcg ctggagctgg 9060
ccaaggcgct ggcccagggc gcccccttcg tcgtgctggc cgagcacctc accatggaga 9120
cggcgtgga gtgctgcgc gccggtgccg ccgctgcct tccgaagctg ctgtccgata 9180
ccacggcgct cagccgtgag ctgagccgcg ccttcaagag agaggtccca tgagactcgc 9240
gtcctgtcc gtctgtctct tcgcggtggc ctgttcctcg gacaagcccc cgccgacgga 9300
gcggccggac gggggcgcg cggaggactc cgggacgggc ggagaggacg cgggcccgc 9360
ggaggacgga ggcacggagc cggatggtgg ctccggctcc gatgcgggca cggatgcggg 9420
agcgggcgac gcgggcacgg tggatggcgg aagctgtgag gcgcccgcg ccgagcccga 9480
tacggagctg gccacgcgac tgaacacgcc ccgccggctc gcggtggaca cgacggacat 9540
ctacatctcc gagtcacact ccctgaacct ccagcagccg agtccgggac cgggacaggt 9600
tctgaggctg ccccgcgcg gaggtccacc cacggttctg gccacgggct tccgcgcacc 9660
agacgccatc gccgtggatg ggacgagcgt ctacgtgctc gacctggacg ggctctggcg 9720
ggtggacaag gcaaccggga agcggggcga cctgcccatac gacgcgacgg atcgcaacgt 9780
caccgtcggg gtatccgagg tgctccgcgc cagctcgcc ggacgcgacg tgctggtggt 9840
cgctacgggg cggcagtcgc tgggtccggg ggacaccgat ggcggaact cccaggttct 9900
ctacaccggt cccggcggct ccctggtgcg tggcgcgga gtggtggact cggacgtgtg 9960
gttctctgctc agcgcgggga cgggagccgg aaacagctcg ggtctgtacc gcgtcccgt 10020
ggacggcagc gcgcccgcg agcggcgga gccaccatac ctccaaggca attcactgga 10080
ggtgaccccc acgcacttcc tcgtaccga gggcgggtggg ggtaccggcc ggggtgtgcg 10140
gctgccccgc gctggaggca ccgccgaggt gctggccgac ggtctccagg gcccggtggt 10200
cccgttgaa ctgaatggga ccatctactt caaggagtca cgtgccgacg gcgcggactt 10260
cctgcgcgc gtgcgcacgt gcgcgctggg gacgtcggat ccggtggggc ccccgggcac 10320
gggaccaggc gggctcatcg tggacggcag caccctctc tacacgtcgc aggagagcg 10380
caccggcggc gccgtgggac gcgtgccctg acgccctcct ccacccccga aaagaccacg 10440
gccccggcg ctcgcgagga ggcacccgg gccgcgtcct gcatcagtgg aacacctcac 10500
cggctcggcg ggcccaggc ctcatcgcc ccggaccgcc gtgagcggct cagtagtcca 10560
tgtcgtcgcc gccgtagtcc ggcacccgg ccgcccgcc gttcttgccc ttcttcttg 10620
cggggcggtc ggcgatcatc gcctcggtgg tcagcagcag cgaggccacg gacgcggcgt 10680
tctgcagcgc ggtgcgctcg accttcgtcg ggtcgatgac gccggccttc tccaggtcct 10740
cgtagacctc ggtgcgggcg ttgtagccga acgcgccctg gccctcgcg accttggtga 10800

tgaccaccgc gccctcgacg ccggcggttg aggcatctt ccgcagcggc tcctggagcg 10860
cgcggcggat gatctccacg ccgaagtcct gtcaccgcc cagcttcagc ttctccagcg 10920
cgggcaggca gcgcaggaag gccacgccgc cgccagggaac gatgccctcc tcgacggccg 10980
cgcgagtgcg atgcagcgcg tcctccacgc gggccttctt ctcttccttc tcggtctcgg 11040
tcgcgggcgcc cacgttgatg acggccacgc cgcccaccag cttggccagc cgctcctgga 11100
gcttctcgcg gtcgtagtcg ctggtgacgg tgtcgatctg cgagcggatg agcttgatgc 11160
ggcctcgcg ctgcctcttg gtgcggcgcg cgtcgacgat ggtgggtgtg tccttgtcca 11220
ccgtgatgcg cttggagcgg ccaggtcgt tgagcgtagc gttctcgtac ttgtggccca 11280
gctcctcgcg gacgaccatg ccgcccgtca gggtgccgat gtccttcagc atctccttgc 11340
ggcggtcacc gaagcccggc gccttcaccg cgccacggt cagcacgccg cggatcttgt 11400
tgaccaccag ggtggccagc gcctcgccct cgatgtcgtc ggcatgatg agcagcggct 11460
tcccggagcg cgccacctgc tcaggatgg gaatcatgtc ctgcatcgac gagaccttct 11520
tctcgtgat gaggatgaag gggtcgtcca tgacgacctc catgcgctcg cggttcgtca 11580
cgaagtacgg agagacgtag ccgcggtcga actgcatgcc ctccaccacg tcgagggttg 11640
tctccaggcc cttggcctcc tcgacggtga tgacgccctc cttgcccacc ttctccatcg 11700
cgtccgcgat gatggagccg atggtctcat caccgttggc ggagatggtg cccacctggg 11760
tgatggcctt cttgtccgcc gtgggcttgg acagcttctt cagctcctcc acgacgacct 11820
cgacggcctt gtcgatgccg cgcttgaggt ccatggggct gtggccggcg gccaccagct 11880
tgaggccctc ctgtagatg gcacgcgcca gcaccgtggc ggtcgtggtg ccgtcgccgg 11940
ccttgtcgga ggtcttcgac gcgacctcct tcaccatctg cgcgcccatg ttctcgaact 12000
tgttgtcgag gtcgatctcc ttggcgacgg tgacgccgtc cttggtgatc gtgggcgaac 12060
caaagctctt ctgatgacc acgttgccgc cttgggggc cagggtcacc gccacggcat 12120
ccgcgagggc ccggacaccg cgaaggatgg cctcacgcgc ggactggtgg aagaaaatct 12180
ccttcgctgc catcgcaacc tcctgaaatt actgggggat ttaactgtga cgacgctggc 12240
cgttagcact cggccatggc gagcgccaac ataaggcca gcgaaatgat gtcaaccagg 12300
aagggcgtgc ctgtgggcga gcggacgggg ccaggggccc tcggggccgt ctggccgggg 12360
ggctactcgg cgaagcgcac gaggttgagc agctgaccca tgtccagggg tttgcgcagc 12420
gtcagcgcca cgcccttgcg caggccccgg tcggtgacgc cctcgtcctg ggcgcccgtc 12480
atcagcagca cgggaatggc ctggaagcgg ggtcggcct tgagcatctc cgtgaagtgg 12540
atgccgtcca ggtggggcat gaggtaatcg gtgatgacca tgcccacctg gttgcgctcc 12600
aggatgtcga gcgcctgcag ggcacccgcc gccagtaga ccgtgtagcc atgcatctcc 12660
aggaagcgag agagcatggt gcacagctcg agatcgtcat cgacgacaag gatgttcacg 12720

8/38

gggcgggccg tcacacatgg acgggcgccc gaacctaaca gccgtgcccc tcgttgacaa 12780
cgcacagcgt gccttcatat gtgccggccc atggggaagc ggaccgcgaa gcgagagggg 12840
gtgggtggcg tggaggctcg gctggatgcg gtggcgagc ctttcgaggc cggggacttc 12900
gaggccgcgc tggccggggc ggagggcctg ctggcgagc cggccgagct tcccagggcc 12960
ctccacttcc gggcttcggc cctgggtgaa ttggggcggc tggaggaggc cgggaaggcc 13020
ttcgggcagg ccctgaaggt ggcgccggag gacctgaaa tctgctgag cgccgcggac 13080
tgcctggtgt gccgcgccgg ggaggaccgg gaggcggtgg cggaggggct cgccctgtgc 13140
gcgcgaggcc ggccctggc ccagaaggcc gacgacgtgg agatgctgta cgagttcctc 13200
ctcctggagg gcatggggct caaccagatg ggcgagtgc ccacggcgct ggtgagcctg 13260
gacgcggcgc tcggccacat gccgcgctcg ctggacgcgc aggtcgagcg gggcatcgcg 13320
ctgttcgagc tgtgccgctt cgacgaggcg aaggccgctt tcgagaaggt gctgaaggac 13380
gcgcgggacg acccgtgggc gcaccactac ctggggctca tcgccgagcg gcgcggggac 13440
gagaaggagg cgaagcggcg cttcgacaag gcgcgggcgc tggtgcccga ggagttcccc 13500
ccgcgggtgg agctggccga ggcgagttc gaccgcgcgg tggaggacgc ggtgaagtcc 13560
ctgccccgcc acgcgaagca gtacctggac aacgtcacca tcgccgtgga ggacctgccc 13620
tcggacgagg atttgctggg gcaggaccgg ccgctgtccc cgagcatcct cggggtgttc 13680
cgcggcacgc ccgtgggcga gcggagcgtg atgaacgcgt atgagctccc cgcgctccatc 13740
gtgctctacc agcgcaacct ggagcgcttc gcgaggacgc gcgaggaact catcgagcag 13800
attggaatca ccgtgatgca cgaggtcggt cacctcatgg ggctggatga ggacgacctg 13860
tggcagcggg ggctggactg acgagcgagg cccgcgtggc ttccagcgtc tggcggatgc 13920
cctcgctgta gggcgtcttg cgcacgctgc ccagcagctt gtgcagcgcc gagtgcgtcca 13980
tgaggacggg ctcggtgagc aggtagtga tctcgaccat ctgcgcatg aacgggttga 14040
agaggccgat gaggcgagc atccccctcc ccgcgctcat gaacttcgcg aggtggccgg 14100
cctgggcgta gatctcctcg acgagcttgc gctgtgaggt gacgcggcg ccggcgaggt 14160
tccaccagcg gccgtaggca ccggggcggt ccatgagggc cgtcacggtg ggaccacgt 14220
cgggcacgta gacgtactcg tgcggcgggt cgatggggcc gatgagctgg gcgcgcttgc 14280
cctgggcggc ggcgacgaag gcggagtga ggaagctctt gtcgatgccg gggccgtaga 14340
agtccgggag gcggaggatg gtggcgcgga gctttccggc ggcgtccgc gcgaggagga 14400
tgtcctcctg ctcccttgcgc atgcggccct tgaaggtgtg cggctcgcg ggggtggtcct 14460
cggtgacggg cgtggtgcg gggcgccgt acgggtacac ggtgccgatg agcaccacgc 14520
gctcgacgcc ctcggcgatg gcggcgctca cgggtcgcg catgagctcc ggggtggagct 14580

ggaagcgcca gtagtccacg ccgaccatgt agatgagcgt gcggatgccc cgggcggcgg 14640
cgcggatggt gtccggcgcg tcgggggttc aggtggcgat ctccgccttc gggtcgcgcg 14700
cgaactgcga ctgaagcgag gcgcgcgagc ggccaaccac ccggtacgcc ctccccctgct 14760
cgcggaagagc cttcacgatg ctgtgtccga tggggcccga cgctccaaac aacgccacct 14820
tgtccgtatc catggtccgc tccttgggtt ccgaacggcg agcgttttat gaacgtcggt 14880
cattgtgaac ggtgttcacg tactgctggc cgttcaggac gtcaagccct acagtgggga 14940
tatggggatt tcggagcgga aggagcgcca gcgggcgag ctgcgcgagc agatccctca 15000
ggtggcgaag gacatggtga cgcgagaggg ctttggtgcc ctctcgatgc gcaagctggc 15060
ggacgcggtg gtagtacgcg cggcgacgct ctacctgcac ttcgagaacc gggacgcgat 15120
tgcccgggag ctgtgcatcc ggggggtcca ggacctgctg gaggcgttcg agccggcggc 15180
atccgtggag gagccggtgg agcggctgta ccggtgggc gaggcctacg tgaagttcgg 15240
gctggagcag ccggagacgt accggctcat cttcatggag gacccgaagc tgtcgacggc 15300
gctgttcagg gacgcgccgg aggacggcg gcggccgcgg tccttcgggg tgctggtgaa 15360
ggtgttcgag gacctgaagg cggcggggag gattgccgag gacgccgagc cctcgaagct 15420
ggccgaggtg ctgtgggccg ggggtgcacgg catcggtggc ctgcggctga cgtgtacggg 15480
cttcaagggc tcgccgccgg aggagctggc gcggctgttg gtgtccacga tgggtgaacgg 15540
attgccgggg ttgaagcccc tgaaggcggc gcggaagaca cggtagagccg ggcgttggtt 15600
gacggggccc gtggggtgcg gcagtaatcg acgggtgtag ggccctccga cctggtggcc 15660
ggcccgtct tcaaaaccgg tgagggcgcg tgagaggcgt ccctggtgag ttcgattctc 15720
atgctctacc gccatgtttt ccttgggggt ttctgacgcg gccctgatcg ggttgacgga 15780
cgccccaccc gcctgtacct ctctgaccca agttttaccc aagtgagggg ggtacgggag 15840
ccgcgcatga cgaagtggag cgggaagtgg atcggaggcc gcacgtacac ggccaccgac 15900
ggaagcacgc gctggatcat ccgcaagacc gtcgccgggg tcgcgtacaa cgtcacgctg 15960
gacgtgcgga gcgaggccga ggccctggct gagctggcgg ccttcgcccg caaccgggcc 16020
gcgtaccgga ccgcgagcca ggagcgccaa gacgaggccc aacaggccca ggaggacgcc 16080
gcgcaggcgg tgttctctgga cgaggatcgt gcccggcggt tcctccagca tctgaagtcc 16140
agcggcaggt ccgacgagta cctgaagagc acgcgcagct acctggccgc gtgggcaacg 16200
gcgctcgcgg acaaggacct gcgcgcggtg accggaccgg cgttgaagaa gatcctggcc 16260
acctgggaca ccggcaagaa gtggaggatc atcgtcttca agtccttcac ctcgttcctt 16320
caggactcgg gcgaactgga cccgatggtg aaccccgac gcttcttgaa ggtccccctt 16380
gcccgccgga accacgaact caagggctac agcatcgagc acgtccagca gctctacgcg 16440
gccctcggct cgcaggccat ccgcgatgtg ctgtgccttc aggccaagac cgggatgcac 16500

10/38

ggcaccgagg tggaccggct cgcgagcggg gacggcaaga tcgcggtgct gaaggaccag 16560
ggcgagatcg ccgccaccgt cacgttcaag cacaagaccg ggaagccgtt caccttgtcg 16620
ctcgatgtgc aggggctcgc cgcgcgccag cgccttcaag cccggggttc agcgccgat 16680
cggcagacca tccgcgaggc cgtggagcgc gtgtgcgcgc gcacgagcct ggagttcgtc 16740
cacttcggcg cgatccggca cagcttcgca acgtggctgg tcgagcaggg cagcatctac 16800
aaccggaagg gcggtggcct gtccctggaa gccgtcgtc aggcgctcaa ccacagcagt 16860
acgeggacca cggccctcca ctacgtcagc gccacagtgc cgccgatgta cgtcgtcccc 16920
atccgcctgg agcattcgaa tgatccatcg gagttgcgca tagctcgggc acctgcgggg 16980
tgaggctcat gaaccagaga gttcggagat cagtacgagg aggttcgttg atgcgtcagg 17040
ccccaggccg gagccttcaa cgttcttatt gcgccgaacg ctccaggtga ttaccgcacg 17100
cctcccacaa aggagccatt cgatgcagat cctggccgta accgaacctc ctagtttcgg 17160
ggggaagtgc accatcaccg acgtggcggg caacgttgctc gggtagctcc acctgcgcgt 17220
gtcctcgatg ctcaacggca gtgcaggagt cgtaaccgat ctccatggcc gcgccctcat 17280
gcaggcgccc ttcaactttg ggggtgtgga tctagaaacg cgacggaggg tcgcatcgcc 17340
gttcgtcgag gcggccttcc cggagactcg cgccgcaatt gccggagcca agccgaccag 17400
agccccatc acggttgcg agttcggagg cgtgctgcgc tcgctgctcc cttccacgcc 17460
cactcacccc gtcgggacta ttcccgctc cattccgatc gacgccatca acatcgttgc 17520
ggcggacctg taccgaacc ggacctgoga tagggagcac ttggaagagt tgacgaaggt 17580
gctcaacacg gagttggagg cgcagtaccg cacgcgcctg gagtggtcag gagacacgct 17640
gcacctgcgc gccaaagtag ccatcgcccc ctcgggcgcg gccagcaaag tctcttcctt 17700
gatcgaggag taggatgcgg cgaactacgc ctcgatcact gccgatcat tcttgtcacg 17760
atgtcggggg gttagagcgt gtcgatccgg ctgcggatgc caggacaaga agttctgcct 17820
cctctatccc gttgggcat gtcgaggggc accaaaacct tcgctcgtga gcttcggccg 17880
cttcgttacg gcgactttca acgaggctca aatcgtgagc ttgcggatgg caacacctgg 17940
gttcggatac gctccggctg cggttccctc ctgcttccca cgccgctcct tgggtgcggac 18000
tcgttcaggc tgcttggttt gttgaggccg attgcctccc ttcgagggg cgccctccgc 18060
gagcagtgcc tcgatgctcg ccaccgtgat catcacggac cgtccgatct tcggtgcccc 18120
ctccaaccgg ctttgcctca ggagcttgaa gacctgtgac cgcttgacgc cgagcattgc 18180
cgacgcgcgc tccagcatca cagccgaggc gttctgtgtg accgattgag cgcaagcgc 18240
acgaacctcg cgcttgatct cggccagttc ctggaggatg gcgcattcca cgggtgcggg 18300
ctcgtggtcc tcaatgggca cggcaggctg aaagactgga atgggaatag gaatgaggtc 18360

caacggtggc tccggccgag caagcaagga gcctgtacgg atatagctga agttctcgat 18420
catggacttg tactcaaccg gtctccgtcg gaggccgtgg atggtgtcga aggtgtggat 18480
cgagccggtg gccaccagtc cgtccctatt ctgaatgtat cgaatctctc cgaggtcctt 18540
ccgggcgctc gccctgcccg cctccgtgcc ggtgaaggcg aacgggaggc gcttggtccc 18600
gcacctgagc cccgttggac gccatggatg gcccggaagg gtcacagcac cggctacttc 18660
agccctttac tcgtgctgog gcgtcatcgt cctgacttcc ccatagcctt ccatcccttc 18720
cgatccctcc acctctctct gcaaccaggg ggagaccgaa gcagacagca tgggaaggcat 18780
cgccgcagtc ggaacgcata cgttccaatc acgggtgcgc aagggtccatc cgagcgccag 18840
cggcgatgct ctccatcacc ccgtcgatct cggctctcca gcgtttcacc gtgaccgggt 18900
acttgattcc atcgctccgc ttccagcccg tgttctcttc accggccaac tccttcatgc 18960
gttcaccgaa cttegtgagc gagaacgcca tccggtcacg gccctggtac ttcaagtcct 19020
cgacgtagtg gccgtacagc gccgtggcct tgatccactg gtcccgtgc gcgcaccggc 19080
gctcgaagaa atcttccacc gagtccttgg ctogaatctc atcctggcgg gcctggacct 19140
gcgccaattg gccgaggagc gggctctccg ccttctcgtc cacgctcgtc cacatggcca 19200
ggaagtcgat ctgctccagt gccgtccagt tcatctgggg cagacagtcc agttgataga 19260
agcgtcgcat tcccgtggga tcgacgacga tctccttcac ctgcttggtg gacgcaccga 19320
tgaagggtggc gttgttggtg ccggtctgga gcttctggac accgagcacc cgggtactcga 19380
tccgctcggc ggtgatgcgg ttcttgagca cgtcagacgc taccgctct gccttcgcca 19440
ttctgtcgaa gaacatgccg tagcagcgag tcaaccggaa ctgctggcgc tcatcggtca 19500
ggaactggag atcggcgggg aaccaacga gatcgccaag aggctgtagc aggctcttga 19560
tggccgtagt cttcccgcgc cgctgtttgc cgacgaagat cggcatcagg tggtccttga 19620
ccggcttccc gaacaacttc cgcttcacct gccaaatgaa gtgaagaacg accacgagat 19680
ctaccgggtc aaggcgctcc gtggttgctc gaatgaactc ctcggcgtca cctcggcgtc 19740
cgtcctcgcg gaacttcaac ttctcacgat actgctcgac gatcccccg cggttggtccg 19800
cctccacag gcgcacggcg ttaccgagat ttgcacgct gatgcatgtg cgctcgttcg 19860
tggaactgat ctcgatctgg gacaagaact cggacgcagg aaacatccgg cccccaatct 19920
ggaagccatc gctcgcgtac acgaatgtga cgccgttctt ctccatccag tgttgaacct 19980
atgtgcgcgg ctgaagcgg ttctcgacct tcggcaacgc ctgcaccaat cggtcgatga 20040
cctcgggctt gcagaccttc ttctcttca aggtcccagc ccacttcttg aagatgccat 20100
ggtcgggtga gaaccaaacc gagcaccaga cgagagcgcg cttgatctcg atctccttca 20160
ccgacccgtg cgctcatgc tgttccacga gccagtcaca gagtgccgaa ggaagcgccg 20220
ggcgggtgga ggccaagtac tgcacgatca ggaactgacg agcccatgct tccgtcttgt 20280

12/38

ggtgatgagc ggcaagcaac ttcttcatct cgcgatgcac ccaatcatcc acaccggatt 20340
tctcgtgctg tgggtcgtgg tccacgtacg gaatctgagg gtcacatccg cgcgcagaaa 20400
ggcccttcgc gatccgcaca acttcacgaa cgacgttcat gttcgtgccg tccatgtccg 20460
gcgcacgcgc gcagacgaac accgtgcggc cctgtgcaag gtatgatgcg gcctcgggggt 20520
gcggtacca ctcgtccgtc aggtcatccc agcttcgtcg ccgggcctca acatcatggg 20580
ccatgctgat gccggggagc cccagcgcgc ggggtgccttc ctgcgcgagc agcgcgcct 20640
tcttctcgcc ctcaatgagc gccatgggga cttcttcgtc ttgaagcatc gggcagagga 20700
tcgcggtgtc gaagatctgg ggcgcgtaac cgaaaggagc ctcgtacttc ttcaccttcg 20760
ctggctcccc ttcgacctca tcctgggtgt acgggcgcag gatgtgggtcc gtgaagggtcg 20820
cggttgggag ggtgagctgc gcctccttcg gccggagcac gcgcgggcga tgcggcttga 20880
ccttggcgta gttgagcccc gggtacacga gcaccagccc tccagcgtcc cccttccatt 20940
cctcgggcca gttcaacaac ttccggttgt tctcgtccgt gtccgttcgc caaccgagtt 21000
tccgcaccgt cgagggcagg aggcaggact tcttcaggtc ctcgcagtct cgcgcgcaga 21060
tcggaagctc gttccactca gcctcgttgc ggctggattc aacctcgttc gccttcgagt 21120
cgggcagtgc cttgtccgag tcctcatgtc ccattgggta cacctctccc atcaacgcgc 21180
aagccgtgct tcgacgcctt gaaatgcagg gtggctcttg aacttgatc gattctggct 21240
acgtcgttgc ggcgccctgg agcggagccg ggttcggtgt ggctacaagg aatatatcga 21300
atgtgtctgc gcacctcgga cgggtgcgcc cgagggccac gggcatcaac gccgctgcgc 21360
gatcctgcgc cgtggacggc caggggattc gcgacgtaca ccggacggca accttcctgtg 21420
tatgagccga ggcaccgcct ggagcacgga gtacctggac gcgatgctcg cggcctggga 21480
gaggtccggg ctgacctcgg agcagttcac ggccacggcc gcgccagggc tgcctgcgcg 21540
taccctacgg ctgcaccgcc agcgaaga gaaggagatc cacgcagagc ttcagcggct 21600
tcgggatgaa aacgcgacgc tccggtcgat gtcgaaagg agcgcgttgg tccctggtgg 21660
ttgccgggcg gcaaccgtca gcgaccccaa cggccacggc attggagatc gccatgtggc 21720
aaacggaacc gaagcccctt cgccgaccgc gaccggcccg ccagcattgg agccgttaca 21780
gccgcgcgtg gtgccgcagg accaggttgc cggaccctcc cccggtgatc acgatggctc 21840
tgtgctcgta ggggaaatca ccgagccggt gccccgaag agaccggcgt tctcctggga 21900
ttgatggcct cgcgcctgct ggccgatgtc ggacgccatc taccgctggg cttggtccgc 21960
gagcgccttc aagagcttct tgatctcggc cgaggacatc gtgcggatgc ggtccacgag 22020
gcgttgccgc tcgtcatcct cgggaaacgt caggagatcc agcggcaaca ctcccaggtg 22080
gtccgccagc accttcaacg tgcgaaggcc aagctcctgc ggccagtgt tagaggtggg 22140

actcagggca cggagcagag gcagagcgcg agcaccaggc aggcgctccg ccgcgctcac 22200
cctgccgagc cctgtggtcg cttgtaatac cctttgggtt tatgaatagt agacctcgga 22260
ccgcgagagc caccggcccc gtccccctca gccgaggccc caccagggag ccggaaggcc 22320
cgagccagtg cccgccgcgc gtcccttccc acctggaggc cgcacatgac cgatgacagc 22380
acgccgaccg acgaaacgca gatcgaggcc ccgacgcccg agaacacctc cgacctgaag 22440
ggcgcccgc cccgactccg gttcaccatc gagaccacca cgtacgagac cctccgcacg 22500
ctggagcggc ggggtaacgg cgagggtgtg ctcttgccg agcgtcacct ccccatgga 22560
ctcgccgggc tcgtcaccat caagcgctg cgcaatcccg tcactttcga gcgctgccag 22620
cggctgatcg aggaggtcca gctctccttc cgctccacc acccggccat cgcccaggtc 22680
catcacctga aaatccatgc cgaccggccg cagtcacg cgagtagct ggacggcccc 22740
acgctggaca ccatcatcag cctcgccacc atgcgcgaga agccgctctc cgcgcccttc 22800
gccctctaca tcgccgccga ggtggccgac gccctccacc acgcgcacac gctgagagat 22860
tcggagaacc ggccgctggg catcatccac cgcgacgtgg cgcgccgga catccgcgtg 22920
gccgcagcg gcgaggtgaa ggtgacggac ttccggcgcta cctactcgct catggtgggc 22980
agggaggaga cgccgggcct cctgctcaag ggcgacgtgg cgtacgcctc gcccgagtac 23040
ctgaaccgta agcccatgga tggccggtcc gacatcttct cgctgggcct cgtcctcatg 23100
gagatgctga cgtgcaagca cctcttcgac gtggaggacg aaaaggcccc caacgccacc 23160
gtggacgtga agacggagga gacgccctcg gtgcccctca cgcagatgat cgcgctcgtc 23220
aaccgctacc gccccgagga cgtggagaac gcgatggcgg gcctgccgga cgcgctcaag 23280
gccatcgtcc acaaggccct ccagcgaaag ctctccgagc gctacgccac ggccgccgag 23340
atgcgcgatg cgctgcgagc ggcgctcgcg gcggagtcac agcccttcgg ccggaaggag 23400
gcgagcgagg agctggcgcg gatgctgtcg gaggcgtccg tcctgcgcga ccgggtggaa 23460
ctggacgaag aaggcatctt ccccgagggc ctggacgccg acgagccgac gcccgaccg 23520
aacgaagagt gagaacgggc ggcggtgcat tccttgaagc aagcgggccc cgggtgtggc 23580
ccgcgtcgtt tcagcgcaac aacacctggg ccaccttcac gagggcacg agctggtcct 23640
cgtccatctt gcgcgccagt ccggtgagct ggcgagctt gggcgcccgc cctgcccctc 23700
gtccccctt gcccgacttc gtccccctc ccgcgtccgc gatccccagc aactcgtccg 23760
aggagatgcg cagcaccgag cacatccgca gcagcgtctg gacgctgggc agcatcttcc 23820
cgcgctccag gcggctgtag accatgtgcg cgaggccag cttctcggcc acctccgcct 23880
gctgagccc gagctgcgtc cgggcctcac gggcggcact tccaatacgg gtcgccagct 23940
cttcattcat gcgtcggggt accaggacac cgagaaccag gtttggcgcg ggcaggaagg 24000
cccaagcctc ggggacgccc ctctctgccg catctcggcg gccatgtcga tccacgagca 24060

14/38

gggccccgtc acgcgggcgcc cctattggcc ggacccatcg ggggcgggaa cctccagggc 24120
caggccgccc aggatgaagt ccacggggac acggaggaca ccgcacaggc ccaccagcgt 24180
ggggaggctg ggcagcaacc ggccacgctc caggcggctg aacaccagcg tcggcacgtg 24240
gatggcctcg gccacctggg cctgcgacag gcccgcccga tgccgggcg cgcgacatt 24300
ggttccgatc tgctgggaga gcttcgctga catgaggaga cctgggagaa agaggagaat 24360
gccgcgtggg cggcactccg agcatacagc tcaccaggtt caacgatagt atccgctcaa 24420
ggtttaccct tctcccgcgc cacttcgctg ggggaactcg tacgagatcc gctaccggtt 24480
gagagcatgt ccaggaacgc cgaccccaca cgccccgcg aaggagagga cacttgagcc 24540
cgctctgca tcccgctcgc cccccgggca cggacatcg cggttacctg gtggaggaga 24600
agctgggagc cgggggcttc ggtgccgtgt accgcgccc gctggagggg cggctctacg 24660
cgctcaagct catccactc tgggggctgg ctgagtggc cgagcgcgaa gtggccatcc 24720
tcctccggct caagcactcc aatctggtgc gcatccgtgg gcatggacag tggccggatg 24780
aggcgctca atccttcttc atcgtcatgg actacgtgga agggcgccgg ttggacgtat 24840
gggccagaga ggagaacccc tcggcccag aagtcgtgct caaggtgcgt ggcgtggcgc 24900
gcgggctggg cgccgcgcac cgggcgaagg tgggtgcatc agacctgaag gagagcaacg 24960
tcatcgagcg cacctccgac ggagaggcag tgggtggtgga cttcggcgcg ggcgggtacg 25020
agagcgcccc cagcatcacc ggcggcgtgc tgccaccggg cagccggag taccgcgcgc 25080
ccgaggcctg gcgcttcag caggagcacg gggacgagcg tggccactcc taccagcccg 25140
gcccctcgga tgacctgtac tcgctgggcg tcgttctcta ttggctcctg acgggcaggc 25200
agcccttct cccggacgag gccgcagggg tggaggcctg gctcaaccgc gccccaaac 25260
caccocatgt gctcaatcca cgctccccg aggcctgag tgccgtgtgc atgcgcctgc 25320
tggccaagga gcccgaggag cgccaccgg acgcgacgc gctgtgcgcg gagctggagg 25380
ccctgctggc ccaggcggac gagtccctgg acgtgaaact ctgcgacgcc tatggcccgg 25440
acacogccac cacgctcgcg gcggtgccg acgcgggca agacgagctg gtgcaatggc 25500
tgaagaagcg caaggcccgg cctcgtcggg ggccacgtcc gccacacggc gagggcgcac 25560
acgagccgga ctcaaacgcg gtggccatcc aggcagagct gccaccccca gttcctgctc 25620
gccccggccc catgccctcg ccgccccgg cacaagccgg tcaattcaag gccctgcgtg 25680
tggcgacctg gatgggtctg ctgggtggtc tgagcgcggg cgcgctggtg atcgcgcgca 25740
ccttggcgcc tccccctgct cctgccgcac gtcacgccct ggagcctcac ccagtgatg 25800
ccgtggactc accacctgca tccccgagt tcgacctgc gccctgggccc cccggccagg 25860
aagtggcggc gccatggatg gcaccggaag ctcaaggaag cgagccgcg ctcgatagcg 25920

cgccccacct agcgggccgtc gccctccccg cgacgatctc cgaggagaag gcttccgtga 25980
agatgaagaa gaacaccggc gtattctccg agcccagacc gcagcgcttc cggggcgga 26040
ccgttgga caa ggcgctcggc ctccgtgtcg cgggctgcct agcgatggcc tgccttgcc 26100
cccaggtccg cccgacgcct ccgccagagg cgtgcccgcc aggtgctgtc gaggccatga 26160
gcaggcttcg aattgagtat gacgaggatg tggtagccac cttcttcgcg accgccagg 26220
agggggtag caggaacatc acggtgcgcg cgggcccgc aacgggtcac ctggggagac 26280
cattgggaaa tctaccggcc cggaccacgc tctccggacg cctcatcatc ggagaggatc 26340
gcgtttacgg ccgactaact gaagcacgca cgccaaaggg tgaccgcttc cctgtatgca 26400
tgcagatcat tgagaacgac tctattgggt tcaagaagga gatcaacgac ggctccaaca 26460
gcgcctcgt ctatccctat tttgacgtga aggcggtccg ccgcttcgag tgaaccccga 26520
ccaggagcaa catggggctg aaccaagtat caatgcaggg ttccactctg ctacttcaag 26580
gtctccgcg atgcccatgt caccatcgt tgttctcctg gtgatagcg tttctccg 26640
ggcagccgc gcgcaatcg gtccagcagc caggggctt ggggtacgcc ggattgagtt 26700
ctcggccgag aacttcgagg caatgacgac accggaggt cagataagcc cgggggtttc 26760
caccacgttc gagttcaact ccgccctcct tcaggagaag gtcgcggttg aaggcgcca 26820
ccgcttctcg ctctgggaca ttggacgaag cactctccg ctggtgccgt cggagcaagt 26880
cctgcccggc gaacgcctgc gagtgacgg gtggttctc gacggagcag cgccggtagg 26940
catggcattc atcctcgtcg cgcaccccg tcaggcgag cgcttggtcg aggtgtggcg 27000
caaccagcg acggtggagt cctatcaaca ggaatcgaaa gaagcccag cggaggcca 27060
aagggtgtac gaagagaac agcgccctcg cgccgagtat gcgggaccgg gcggcctcg 27120
tggcctcctt gcgaaccg tgatcggcaa gtcgggcgtt tcggtcaagc ccctcgactt 27180
cgagaaggaa gtccgccagc gtcctggaga cgcgatcagg attcttcgt catggagcta 27240
ccgctccgc aaccgctgt ccgtgctcat ggatttgac taccagaag ccgcgcaacg 27300
ctggatcgc caggagcat cgttggtgag caagacagg gaaacgctga gcatgctccc 27360
tgtatggcag gaagcgccc ttgccaatgc ccgatccg cgctcgtg tggaagcaga 27420
ggctgagccg gacgcgccc aggacagtt caccctcaaa ctgtgggaat cgaacgggt 27480
gcgcacgat accctatccg gggtgacatt ccataatgt tgagaacatc ctactatca 27540
agtagcacc cacatagcat gaagtcgcg ccggtgcgtg tcatgggccg ttcaaacc 27600
tgaaaaccaa taactacatc cttacacat ccaaaccaag ccccatgtct acagccgagc 27660
aaacgaagaa aataatcaag aagctgcgag ccaccta aa cgaagggaagc tcagaaaaga 27720
ttcaagcact catccagtcc atcagactgg agaatccatc tcagtggctg aatttgatcg 27780
ccctcgact gggcccata gatcctagga atccaacatt ctcgaaaaa gtatccgca 27840

16/38

taaacagatt ggcagaagat cttgcgccag aagagcagca cagctacgc ggatacga 27900
gcatcgcaag ggccatgcag atcgtatcgc aagagtccga caaatcagaa ttctggaaat 27960
cttcgcgcct cacaatgatg tggcgcctat ttacgtacgt tgaatccaa tacatatattg 28020
ttagcaatac gattcgaaag aagctcaagc gggccccacc gggctcgcca tttgtcatgg 28080
cgagcccttc aatcgcggtt ccaggaaccg atgggaccgc aacgacaagt gtcgatcttg 28140
tttttgaatc taccgccgaa gccgccgaac taatactccg catctccatg cacaggatcg 28200
acaccaacc cccaaatgaa aagtttacct ccagacccc ttatgacgat cctgactttg 28260
gcaggctcct cgcttgccga cagcgtggc atcgattcag cctgctatgg aaccaaattcc 28320
agtataccga ctggcgacct gagcatgac cagcggctt catctggaca ccaaaaaaca 28380
aggcgcaata cctccaaagc gaaatcgga acattcgga gcgcctat ttgcagcaag 28440
caatgctgtc tgatggactc ggtcctgggg cagaactgga gagtagaaga gaagtccctgg 28500
aacgcctcaa gaggagcttc aatctccga agccaggga aacatgggac cttaagattg 28560
atattaaagc cctgcgagat ctctcggtc ttgatcgata tgccacggct cgcaagatct 28620
acatgtcgcg tcgattccta accccagtaa tgggagcgcc agcagatcct gaactggaga 28680
agtcacagaa ggagttgaag ttgaacttca cgctgcatga actcgccatg gctctgctgc 28740
aaaggaatca cgcaacattc aagaccaagt cagcaatata caacattggt tttgcagtca 28800
ccaagcagcg cctatgcaag gtaatcagcg aggcgtcggg catcgaattc gaattctgct 28860
tttcccatat caacaacatc gtcttcgacg cacagagaaa aaacctagag atctgggacc 28920
agccgctcgt accagtagat gaccataggg tgttggtcgt tccttcgatc attgccaatt 28980
cccctcctat tcgcacaata gagcatttag cttcagctcg ctcaagctcc gcgtttgact 29040
cccggggaga gtattatgag cagatgattc aatcgaggct tgaggaatta acacaagcag 29100
tcgctgtgcg gggcatcaaa tttcctgctt cagatgggcg gcagatcgaa tacgacttgg 29160
tgctttattg gcaagggcac ctatttatct tcgagcttaa gtgcctgaag atgacaacaa 29220
ctcctgtgga tgagttcggg gcgagggagg agatcgagcg cgccgagcac caattggtac 29280
ggcgtgtttc tattgccgaa actgactgga aggagttcag gagccaagca caggagttt 29340
cactccccga gactccgcca ggcgcagaa acatccatgc cataatactg accaacatca 29400
tgaactttac gggccgaaga actggagaga tcgtgacggc agacgactcc tgctgtttta 29460
gatacttcga cagccctcat gttcccgat ttgatgctgc cacgggtgcc gaaatctcaa 29520
gcattgcacg tttgcgcctt cgcaacgagg ggccgccatc tgcttcgggg ctttggaat 29580
atttatcaag ccctccgag ttggagtgga tgcggaagca aattcggcct gaataccacc 29640
cactgccacc gctggatgag ggcaactatc actggaagct catggtcgca ttcgatccgg 29700

ccgacgcccga gacgactaaa tagtggttggt ttggtctcgt tgcgaagggtg agcattgacc 29760
tcggctggac gctggattca gcgtctacaa caacgatatt gcagggtgag gcgagggtctg 29820
cgtagccttc gtttagaacc cgtggagaac ccgacccatg acgctatacg aactgtttga 29880
agccgatgat gaccgcgatg tccccgccga ctttccttgg ctcttgatcg agatccctga 29940
ggagcaactc cgccatgcgg cgcttctcct gaatgagggt gagtggcgac cctcgcacat 30000
ttctggcatc tggtatcggg tggatcccga gcgcccagct cagaagcagc agcggcacgt 30060
ccacgtcgcc gcgaagaagc acatcaagat ccccacgaag caggcttcat ggaacagaga 30120
caccacccgc catgaccgca agacattcaa tgcgaagctg ggctcgcaag ggagctatca 30180
agacgtggcg aagacggcg tgggcctgcc gcccgaggct gtctctgaac acgttgttga 30240
aaatccggcc gaacaacaga tgctgctcac cgaatccact gatccccagt tgaaaactga 30300
gttctacagg tggttggagg ccagcccag gcggtctaac ttcgaagcgc tcgtcgagga 30360
cgcggttagg ctcaatctgc cgctcgacct ggaatcgtaa ggagtggtt aaatgcggtg 30420
atacgagcgg tgatacgagc gcctaacaaa actggcatag ggacccggct tctcaaggac 30480
gcagctgcaa ccaaggcgat tttgtaggca ttctttgggg ccgagacag cgcgcgggat 30540
ctccggcacc cagttcacgg acgtcacgt ccccggtgca gccaatggcc cctgaccct 30600
gccgtcccga ccctcgtacc caagctgtac ccaagatgaa cgacgcggg tggacctcca 30660
tggacgtcc acccaccgcg ctccagagg aaaccggccg accctcaatc cctccgggca 30720
cctcgcggtat cgggggagcg gggcttcaaa accggtgagg ggcgctgaga ggcgtccctg 30780
gtgagttcga ttctcatggg gcgacgtgaa cccggccgtg ctccatgcgc tccacaacgc 30840
ggcggaagat gaggtcgata atggcgagg actgagcctt caggagtggc cggaacagac 30900
cagggccacg tcagaagcga acgacctccc ttctgtaacc gaataacgaa acagaggcga 30960
ggatcacaac ttctgaaatc cgttgacgcc cgtgggcttc atttcaggaa tcttgggaaa 31020
tgctctcct tgctcttctt gggcttctcc atggtagcgc ttcgcgcacc caggctctgg 31080
tactatccat catatggagc gctccctcgg aacgacagag caggctctct ataaggtacg 31140
gcagaagcgc cccttcaaca tcgtcattac gggcaagctg cgaggacgcg tggaggagca 31200
ggctgtgcgc caggggttg tggccctgca acggcgacac ccgtgtctcc gagcgaagat 31260
cgccgcgggt actccgccac gcttcacctc gcaagggtgtg cccccattg cgctggaagt 31320
ggtggaacgc aaggcagagg agtcttgga acagcatgtc gaggcagagc ttgacaggcc 31380
tttcaactcg gacctgggc cgctcgtccg cttcatcctc gtgcggggcg agcaccgctc 31440
cgaggtgctg tgcagctatg atcatctgat cggagacgcg cactccggca tcttcgccct 31500
ggcgagatctg ctgcaggta tggcctccca ggaccagcac ctgcccagc tcgccccaa 31560
ggcgagatctg gaggagctga ttgggcccga ggtaccgggc acccgactgc ttggcgccgc 31620

18/38

cgtgcgccgg ggctcctcgg cgctgctgcg cttgtctccg accctcaata cgctcacaga 31680
gcggctcgtc gcaccgcggg gcatggctcg cagcagcgcc gagggacagg cgctctacag 31740
ccaccgaatt ctggagcccg agcagctcgc gcggctgctt gcgcgctgtc gagagcaagg 31800
gtcctccgtc cagcagcgc tcggggcgcg gctgctgata gctcgggccg agtcgcaggg 31860
tgccaagaag cgggtgacgc tcacctcac ctctgcgttg gatgctcggg agcgcttttg 31920
cgtgggtgag gacttcggcc tgttcaccac ggggaagacg gacttcttcc gcgcgcgcgg 31980
cagcacccca ttctgggagc tggcgcacag actccgatcg cctctgcagg cagcacggaa 32040
gaagcggagc cacttgaggc tgttcgcct gatcctggag atatccgcgc tcaccgtgga 32100
catctcctcc cagccctgga tggaaacgcg gacacggctg ggctgcact ccatgctcgc 32160
cctgagcaac gtggggcgtg tggagattgc cgcccgtat gggaacctga cgctcgaggg 32220
ccttggttc acggggacga ccgcggccca gttcgacgtg gtgctgacgg cgatcacctt 32280
tgcccagcgg ttggaggcca gctttctctt caacaccgcg cacatgcagc gcgagcaggt 32340
ggagcagctc gcctcgcga cctgggagct gctggccgag gcgacgcgt gagcggctac 32400
tccggtgga tgtgtacat ccgcttgacc tcttcacgcg gcagctccat gaccttcag 32460
ggatcccagt ctccgaagag gtccacgctc atcgcgtgc ctgcctccac cgcacgccgc 32520
agctggggaa tgggaagcgc ctccggcagc aggcgaacgg gcaccaagga cttgtactgg 32580
gccgatcat agttgcgggt gcgggggccg aacatgaagc ccaggctgaa ggccgtggtc 32640
agcatttctt cccagatgga gatgccatag ccggagagca ggtggacgaa gtcatgcacg 32700
gccgagtga cgggcagggc gtgaggggtga cccggcagcg gcacgccatg ggagtggcag 32760
tacttccaga actcgcggcc aaaggaacct tcgggcaggt gctcgagctg ctgaaagcgt 32820
tgcttgaggt ctggcgctc attgcgctcg cgcagcatgt catagagctt ccagggcagc 32880
tcgcccttcg gcagcccaa catgcggcag aagtactcaa tataggcagg cacgaggaag 32940
gcggtcagtc caacggacag ccacagcttg gagcgcgcgt ccgcggccag ggaacgctcg 33000
cgcggtccag gggctcgtg gtacttccgc agaaatacat ccgtgagtat tctcattcga 33060
acaccttggc tcgttatacg gctggggcag ctttcaggca ggctcgtgtc tcctggagaa 33120
agccatacgg atccttgggg tcctcgcca tggactccac ctggagcagg cggcctaccc 33180
ccagccgctc ggcgtgcgcc aggatgactt ccagcaccac cttgtcctcc aaggaaaacc 33240
cggtgagtc gaagacggtg aggcatcgc ggtgagagcg ccactgctcg ggagagcggg 33300
tcagttcaaa gaggtgggg ccaatctctc caggctcgag ctgctggcac tccccctcca 33360
ccacggcctg ggggagatag tccgggcata ccagggcgcg acgcagcatc tccaggggca 33420
gctctgtctt gccgggcaga tctgagccca cggcgttgat atggaggtgg gccttgatca 33480

gacgcccgga caccaccggc cccgcgcccga cggcgacgga ggtggccgtg cagaggatgt 33540
ctgcctcctg ctccacgcgc gtcaggggga ccgccaccac ctccaggccc aggaaggcca 33600
cacgcccggc aaagcttcgc atggcttcgg gagcaatgtc gtgcaccagc acccgtccca 33660
gaggaaagac gcgcgagagc gcatgtgcct gggtgacggc ctgagcgcca caccctacca 33720
ggccaatac ccggtgtcc tcccgcgcga gcagccgact cgccacggcg gagggcgccc 33780
cggtgcgtag cgcggttaga aaggtgccat cacaaacggc ccgcaggtgc ccggtcgtca 33840
catcatacag actgaggag gaaagaatgg tgggcaagcc cagctcctcc gggctgcggg 33900
ggtttagacc caccagcttg atcacggcgt ggcccccac ctgcatgacg ggcatccact 33960
ccaacacgcc ggggtggggg tgctgatact ggaagccctc gcgcttgccg agctcgggtg 34020
tggcggggtc gaagtgtcc agctcgtact cgagggtgga gatgacctca tccatcagcg 34080
cgtgcagccc cgttgacagg acgatctggc gcagttcctg ctcaccagc atccttgtca 34140
tccgatggc tgaccccgcc cctttcagct gcatgacagg cctcactgca ttcacggctc 34200
cagttcaggg ccgggcaaaa tctatcttca acatcccctt gtccaggacc gcgtccacgg 34260
catccagctg ccgccaccgag cgcggcagca tgacgtggcg gcggaagccg ccacgcgga 34320
tgatgaggtc attgccctgg cgatccagga caatgtctcc gcggtcggca cctggcatga 34380
ggagctgcag ccgatagcgg ccctcgtgct gcttggcgaa ctgatagggc tgcccggaga 34440
tgtagcgctc ggtgggggtcc ggtccccgt agagccgctg ccgcagatcc tccaaccgct 34500
ccagccccac cacctcgtgc tccagcagcg ggagcctggc caccggcatg ggggagaagt 34560
actcctggat gtgatccagg taggcagatt gagacttggt ccactgagca aagtattccg 34620
tgtccggcag cagacgattg accacgacct ggtccaccgt catgccatac atgttgaagt 34680
agaggtaggc ccgctgtgtc tcgcggatga ccattttttc cgcgctggag accaatcgca 34740
ccgtggtgcg attgccatcc gtgagcaacc cttcaatgcc cttgacgcga tcgaagagct 34800
gctcgaggga ggcaaagtag ctgtcctccg gcagatcata gctcgtgagc cgactggcca 34860
ggggccgggc cagcttgacg agcgtgcgat ccacattgaa gcgccggcg atataccact 34920
gcagcacgct ggtgatgttg acgaagcgca gcgactcgcc cgtgggcgga cagtccacga 34980
tgatcacgtc atagcgcgcc tgctgcacgt actggttcag gtagatgagc gagatgacgt 35040
cctccgtgcc cgggagaatg gcgacctcct cggccaccat gttggatacc ccggtggagg 35100
tcagcagcac ggacatgtag ttatagactt cactccactg ccgctggagt tcgtgctgga 35160
tgtcatctc ttgcagttca aggttggggg cgaccttctg gggcagtccc tcgttgaagt 35220
cgaagagctt gcgatcgaga tcgaagctgt ccgagaggct gtgggcaatg tcgaaggaca 35280
tcaccaaggt gcggtagccg cgcttggcag cggccaccgc ggtggcccg gatgccgtcg 35340
tcctgcctac cccacccttg ccagagaaaa gaagaattct tgtcatggat gcccgcgtg 35400

20/38

cctgccggca	acggggccgcc	atggcgctct	ccccctccat	gggtcacctg	ttgcgtctgt	35460
ggaactcatc	actatcagcg	acatgccaga	aggagcaagc	tggacggtgg	cttgacttgg	35520
tatgacgctg	gacagtaggt	atgccaatag	aggtgggggg	ctttatgaaa	tcgatggatt	35580
ggctatcggg	cacagccccc	ccccttggtc	aattccagga	ggggtgcaat	tgacagccac	35640
agcgactgcc	gagcacgcgc	tgaaagccct	ccggggccaag	ggactcgcca	cgtcggacga	35700
ggcgctgccc	gcggtgcggc	acgccagcgg	cgaaaaggct	cccctctcat	cctcacagcg	35760
gcgcatgtgg	ttcctggagc	agctcgagcc	aggcaacgcg	gcccagcatt	tgttgcaaag	35820
tcatcacata	caggggcctt	tgcaggtaga	ggggttgcg	cgggcggttg	ggctcatggt	35880
caagcgccat	gaggcgctcc	ggctgggtgg	tgccattggc	gcggcgggcc	ctgagcagcg	35940
tgtgcggccc	gagtgggtgg	ccgagctgcc	tctcagcgat	ctgagcgggg	tcgcccaggc	36000
ggacagggct	cgggccctgg	cggagctggc	gcgcgtggag	gctgcggcgc	ccttcaacct	36060
gcagcaaggc	ccactgttcc	gagtccggct	ggcgcgactg	gcggaaacgg	agcatgtgct	36120
gctggtcacc	ctgcatcacc	tgatctcgga	tggagcctgg	agctgcgagg	tgctcatcaa	36180
agagctggcg	atgctctacg	gccagcacgt	ggaaggctcg	agcccggagc	tggcgccact	36240
gcccgtccaa	tacagggact	acgctggctg	ggaggcaagc	ttcgcgcccc	cgggagaggc	36300
cttgagggcc	tgggtggcgg	agcggctcgc	cggggtgccc	acggtgctgg	agctgcccac	36360
ggagggagcg	cgtccctca	gacagacgta	ccgggcaggg	cgggtggcca	tcaccgtgca	36420
gccgcgcctg	cgcagggcgc	tggaggacct	ggcgcacaa	gagggcggtg	cactctttgc	36480
cctgctgctc	accgccttca	ccacgctgct	acaccgctac	tctcgccagg	aggaattggt	36540
ggtgggggtg	cccgcctcgc	agcgcccccg	gcccagctg	catggcctca	ttggatactt	36600
cggcaccccc	gtggcattgc	gctcgcggct	ggagcctcac	acccgggtgc	gggaggcgct	36660
gcggcagctc	gatcaggagg	tgctgaggc	caacgcccc	gcggcgctcc	ccttcgagcg	36720
gctggtgagc	ctgctagaca	ttgcccgcag	ccccagccgc	caccgctgt	tccaggtgct	36780
gttcgacctg	ctgccggagc	agccccagcc	taccgctggc	ggctgcgcct	tccggccctg	36840
ggagtccttc	accggcctgg	tggcgatatga	tctcactctg	ctgctggagc	cgcgcggtga	36900
ggggctggag	ggggcgctgg	actacagcgc	ggacctgttc	agcgaggcgc	gcatgcagcg	36960
cgccgccact	cagtacctgc	acctgctgga	gcagctcgtg	gagcggcctc	acgagaggct	37020
ctcgcggtg	gcgctcctca	cgcctgggtga	gcacgaggct	ttgctggcgg	ggcacggcct	37080
gggcgggccc	gtggagggcg	cgcaggtgct	ggcccaccac	cgcttcgagc	accaggtgcg	37140
gctcaccccc	catcaccctg	cgttgtgtct	cggcccgcag	gtgctctcct	acgagcagct	37200
caaccgccgc	gccaaaccgc	tcgcccaccg	gctgcggcgc	ctgggcgcgg	gcccggaac	37260

cctggtaggg ctgtgcgtgg agcgcctccct ggagctgccg gtggcactgc tggccatatg 37320
gaaggccggg gccggcttcc ttccgctgga cgtcaatcag ccccgcgagc ggctcgcctt 37380
cctcctggga gatgcgagct gccgcatacct cctcaccag gagcacctgc tccagcgcct 37440
gccccccacg aacgcgcgcg tectgtgtct ggagagggag gcggaggcac tggagcgcga 37500
gccccaggag gatgcgccac acgaggccgg gctcgacaac ctggcctacg tcatccacac 37560
ctctggctcc acgggcacgc ccaagggcat cgcgatggtg catcgtctgc tggccaacct 37620
ggtggcctgg cagctcacc atgagcggct gggtgcccc tcacgcacgc tccagttcgc 37680
ctcgtcaac ttgcacatct gctaccagga actcttcacc acctgggccg ctgggggcac 37740
cgtggtcatg gtgacggaag aggtgcgcag ggatcctgcc cggctgctgg aagtgtgga 37800
gcaggagcag gtgagccggc tgtacctgcc ctctatgcc ctgcagcagc tcgcgcgggt 37860
ggcggatgag cgcggcgcgg cccctcgcca cctgcgccag ctattaccg ccggggagca 37920
actccaggcg acccccgagt tgcagcggct gctctcgcgg atgccggagt gcaccctcca 37980
caaccagtac ggccctccg agtgccacgt ggtgacgagc catgacctga cacgcgagcc 38040
gagccgttgg ccacggctgc cgcgcgtagg ccggccactg gccacctcc ggggtgctgtt 38100
gctggacggg gagcagcagc tcgtgcctcc ggggtgtggc ggagaggtgt tectgggcgg 38160
ccccgcctg gcgcgcggct acctgggacg tccggagcag acggctgatc gcttcgtgcc 38220
cgatccattc tcccgcgagc ctggagcccg gctctaccgc accggggacc tggcgcgcct 38280
ccgagaggac ggagcgtg agttcctgca acggatggac gcgcaggtga agatccgggg 38340
ctaccgcata gagcctgggg agatcgaggt ggtgctgtgc gagcaccg ccgtgcacca 38400
agcgcacgtg cggccctacg tggacagcgc cggagagcgg cgcctggtgg cctacgtggc 38460
cgacaggctc gaggacacgg acggggcgga gacagagcac gtggagcgcct ggcgcgcggt 38520
gtgggatgag acctatggcg gcccaagcgg caggcccttt gacctggccg ggtggaatga 38580
cagcgtgcgc ggcgagccgc tgcctccgga gcagatgcga gagtgggtgg agacgacagt 38640
ggagcggctc atggagctcg tccccggcg ggtgctggag ctgggctgtg gctccgggct 38700
cctgctccgc cggctcgcgc cccgctgcga gtctactgg ggacaggagc tctccccggt 38760
agccgtcgag cggctgaggg agcaactcca gacgggcggc tccccgctcg cgcagcgcgt 38820
gcggctcatg gccagccgg cgatgactt ctccgggctt ccggaagccg gcttcgacac 38880
cgtcatctc aactcgggtga cgcagctctt cccagcgtg gactacctgc tgcgggtcgt 38940
ggagggcgcg ctgcgcgtgc tccaaccggg cggaaccctg ttcattggcg acgtgcagaa 39000
tctgcggctc ttcgagctgt tccatgcctc ggtggccctg gagcaggcgt cagcggacct 39060
ggagggcgcca gcgctgctcg cccgcacccg ccagcgcag ctgctggacg aacggttgta 39120
cgtggaccg gacttcttcg ccgcgctcgc cagcacttt cccagctgg gcgcgggtccg 39180

22/38

gctgcacctc aagcgcgga gcgaggagaa cgagatgaac cgcttccgct acgacgtgga 39240
gctgcagctc gccccgtag cgaaagcagg ccccgcgcaa gagctgccag agctggactg 39300
gagggcatgag gggctcggcc tggaaaggct ggagcgcag ctggcgagc gaccggcggg 39360
gctggtgctg cgcaatgtgg ccaacgccc caggcgagc gaggcggcac ggctggcgct 39420
gctgcggact ggcaacagcg tgggcccggct gcgggactg ccggcggtga catcctggaa 39480
tccagagcag ctctggaggc tggcgaggc agccgactac acgtgccacg tcacctggag 39540
cgcccaggac gaggaaggcc gtttcgacgc gctgctcatg gcccgcgagg ctggctcctc 39600
gcgcccggcg gcctggctga cgctccccc gcctccccct cgacctgga agagctacgc 39660
caaccagccc ctggcgcca gccgcccggc caccctcgct ggagtgtctg gaagccacct 39720
cgagcacaag ctgcccagat acatggtgct ctcttctctc gtgctgtctg acgcgctgcc 39780
gctcaagccc accggcaagt tggaggtggc ggcactgcct ccccccagc ccgcacaagc 39840
cgagcaaacg ccaggccacc tcgcgccggc gacccccacc gagcgccggc ttgccgagct 39900
gtggcgggcg gtgctcggcg tgcctcgctg gggggtggag gacaacttct tccagctggg 39960
gggcccactc cttctggcca caggctctt gtcgtctatc cgcgagagc tggggctgga 40020
gctgccccctg cgggtgctct tcgagcagc caccctggcg gccatggcg gttgtctcga 40080
ggctcagagc tggagcacc aggtcccca cctccctct gcctccatc agggaggaga 40140
actttgagcg ccggagcgtt gctcgccac gccgagagc tgggtgtacg cctgtgggta 40200
gagggcgagc gactgcgctt ccaggctccc ccagggtga tgacgccga gctgcagtcc 40260
cgctgggag gcgctcgcca tgagctatc gactgctac ggcagctcca gccctcctct 40320
cagggaggca gcctcctggc ccagtggtc cgaaacgggc ggctggcgct ctcttcgcc 40380
cagcagcggc tgtggttcca ggagcagct caccctgagg cccggccaa caacctcacg 40440
ggggccgtg tgttcaccg cccgctgcac gtggccgagc tgctcggggc ggtggtgctg 40500
ctggtgaggc gacacgaagc gctgcgcacc accctgggag agggggcg agtccccat 40560
tcgctcatc gcgagccgtg gcagccggcg ctggaggtg aagcgctacc gggtgccacc 40620
gtggcgagc ggctggagca ggcccagagc gtggccctg ccgagtccc gcgcccgttc 40680
gccctggaga cggagcccca cctgcgctg cgctgctgc gcctggccga gcagcagcac 40740
gtgctggtg tcagcctcca ccacatcgcg gcggatggag tcgggctgca ggtgctcgaa 40800
caggagctg cggcgctcta cggcgcgct agcgtgggg ccgagccgag gctacctccc 40860
ctgcccctgc aggtggcgga ccttgccgac tggcagcgcc ggtgggtgga gggagaggag 40920
taccaggtc agctcgctta ctggcgcgga cagctcgccg ggctcacgcc cctggaggta 40980
ccaggagatc atcctcgccc gcgtatcccc tccatgagag gcgcgagggt acgggcgccc 41040

ctgctctccg ctccgcaagc ccaggtgctg cgcgcgctgg ggcaaggcga gggcgccacc 41100
ctctacatga cgctcctggc cgccctgggc gtgctgctgc aacggtggac gggccagcac 41160
gatatggcgg tgggcagcgc cgccgccaac cgcaaccgcc cgggcttgga gggcatcctc 41220
ggcttcctcc tcaacatcgt cctgctgcgg ctcgacttga gaggccgccc ccgcttccgc 41280
gagttgctgc gtcaggcacg caggggtgtgt gtggaggcct atgcccatca ggagttgccc 41340
ttcgagcacc tggtaggagg gctccagcca ggcagcgagc gcggggacag ctcgctctat 41400
cgcgtagcgc tcgcggtgag cgacacgcca tggatgcccg gccacgggct gaagctggag 41460
ggggtgcaag cgcagccgct cgactttccc cgggggggtgc tggacctgga tctccacctg 41520
tgggtgtatg acaccgggga ggggctcacc gggcggttgg agtacgccgt ggatctgtat 41580
gaggagccca cggcaaggcg gctcctggag ggcttccgcc aggtgctgga ggcggtggtg 41640
gaggcgccgg atcgccccgt gccggagctg ccggtgctgg gagagcagga gcgccaccag 41700
gtgctctccg ggtggaaccg cacacagcgc ccctatccgc gagaggccag cgttcacggc 41760
ctctttcagc agcgtgccct ccaggccccg cgggccgtgg cagtgggtga tggcgagcgc 41820
agcctcacct acggcgagtt ggcggagcgg gcccgcgggc tcgcacaggg tctggtggcc 41880
cgaggcgtgc ggcgaggaga cctggtcgcc cttcggctgg agcgcctgcc ggagcaggtg 41940
gagtccatgc tggccgtact ccaagccggg gccgcctacg tgccgctgga tccttcctac 42000
cccgtacagc gccaggagtt catgctccag gacagcggcg cgcggttgct cgtgcacagc 42060
ggcccgctgc ccttcgctcc ccagggtgc gccacgctgg atctccaggc atggcacccc 42120
gctccctccg acggtggcga gccctgccc cagtgcagcg gagaagatct ggcgtatgtc 42180
atctatacgt ccggctccac cggccagccc aaggcgctgg ccgtgtgcca ccgggccatg 42240
accgcctgg tgtgcaacac ggactacgtt cagctcgggc cagaggatcg ggtcgcccag 42300
gcttccaatg cctcctttga cgcggccacc ttcgaggtct ggggcgcgct gctcaatggg 42360
gcacggctgg tggggctcgc cacggaggag gccattcagg cgcgcggct gcccgaggtc 42420
ctgcgcgagc agcggatctc cgtgctcttc gtcaccaccg ccctcttcaa ccacgtggcc 42480
cgggagcaac ccaggccctt cagcaccctg cgctacctgc tctttggcgg cgaggcagtg 42540
gatgcctcca gtgtgaggcg ggtgctgaaa cagggcgctc caggacacct tctacacgtg 42600
rarggracca cggagaacac cactttctcc accgcctggc ggggtggaaca cctggccgag 42660
caggcgcata ccgtgcccac gggatcatccc atcgccaact cgcggtgca cgtgctggac 42720
gaagcgctgc agccggtgcc ggtgggcgcc atgggcgagg tatacctggg aggtgatggg 42780
atcgcgctgg gctactggcg ccaccccagag gccaccgcgg agcgcttcgt gccggatccg 42840
cacggcctgg agcccgagg ccggctctac cgcacgggag atctggcgcg gcggcaggcg 42900
gatggggccg tgggtttcgc tggccgcgtg gacaggcagg tcaagctgcg aggcttccgc 42960

24/38

gtggagccgg cggagatcga gtcgcacctg tgcgagcact cggaggtgtc cgcggcggtg 43020
gtggagctgc ggggcgaagg ggccctgcgg cggctggtcg cctatgtcgt cccccgcgcc 43080
ggtggccgcc caggtgccga ggagctgcgc accttcctcc gcacccgact gcccgagtac 43140
atgctgcccc ccagcttctc cctgctggag gcaactgcccc tcacccccaa cggcaagggtg 43200
gatcgctcgg cgctaccga gagcttcgag gaggcctccc gcgagcaggc tccggtggta 43260
cctccacgag gccccgtgga ggcgctgctg gtggacatct ggaggagggt gctcggcacc 43320
caacgggtat ccgtgcacga cgacttcttc gacctgggcg gacactccct gctggccacg 43380
cgggttgtct cccgcttgcg agaggcgctg caggtggagt tgccgctgcg caccctcttc 43440
gaggccctc agctctcggc gctggcgggc caggtggagg tcctgctcgg ccaccggcag 43500
ctccgcccc ctccgctcgt gcccgtgta cgtccaccag agctgcccct ctccctcgc 43560
cagcagcggc tgtggttctt ccagcagctc gcgccacaaa gcaccgccta ccagatcctg 43620
gatgcgtggc atgtgcgggg ccgggtggac gtggcgccc tggagcgcgc cctggagcag 43680
ctcgtgcgcc ggcagtaggc gctgcgcacc actttcgagc ctggcggcga cggcgtgcct 43740
cggcagcga tccacgcgcc ggctccggtc ccgctgcggc aggtggacct gcgctcccat 43800
ggcatcgccg cgcgggagga ggccttgcgg tggatgcgcg agcaggcgct ccggccactg 43860
gagctggaca aggggcccgt gctgcgcgtg tcctgctgc gcttgaaga ggcacagtcc 43920
ctgctcttcc tcgagctgca ccatatcgtg ggcgatggct ggtcgtgag cgtttggagc 43980
cgagagctgt cgcacctgta cgaggcggcg ctccatggcg ctgagccggg gttggcgccg 44040
ctgcccgtgc agtacgcgga cttcgcgtg tggcagcgcg gctggttgca agggccggtg 44100
ctgcgggagg agctgacctg gtggcgggag cggctcgcgc gcctggcccc gctgcgcctg 44160
cccgcggacc atgcccggcc ggaagtgcag cgcttcaacg gcgcgacgta ccgcttcacg 44220
ctcccgggcg tacgggtgca ggcgctgcgg cggctcgggc atgagcacgg ggccacactc 44280
ttcatggtgt tgctcgcgg cttcaatgcg ctgctggcgc gctacaccgg gcagacggac 44340
atcgccattg gcgcgccat cgccaaccgc acacgtgggg aggtggaggg cctcatcggc 44400
ttcttcgtca acaccctggt gctgcgcacc cgctggaag gcaaccctc ctccctcag 44460
ctgttgaggc gggtaggga gacaacgcta gaggcctacg cccaccagga gctgcctttc 44520
gagaggctcg tcgaggagtt gcagcccag cgccaggcca accagaacct gctggtgcag 44580
gtgctgctcg ccctgcagaa cgcgccccgc gagccgttgc ggtggcggg cctggaggcc 44640
gagcacctgg agtacctggt ggccaccacg cgcttcgacc tcgagctgca cctgtgggag 44700
gaggaagagg ggctgagctg catcgccgtg tatgaccgcg cgctgtacgg cgcggggacg 44760
gtggagcggc tggtagggcg ctggtgcacg ctgctggagg gcgtggccga gctcccggcg 44820

cgccgcgtgg cggagctgcc cctggtgcct gcccgggagc tgcgacaggt tccgccgccc 44880
tccccgcgcg cggcgatcga aacaagcatc ggcgcccgt tctcggaggt ggcgcggcgg 44940
cagccaggcg ccacggccgt caccgaaggc gggcggcacc tcacctatgc agagctggag 45000
gagcgctcgg aacgcctggc ccgctacctg gcctggctgg gcgtgcgcgc gggcgacaga 45060
gtgggggttg ccaccgagcg cacgctggag cggatcatca gcctgctagg catcctcaag 45120
gcgggagccg cctacgtacc gctggatgtg cgccagccag cgcgtcggct gagcctgctg 45180
gtgcaggccg cgggcgtgcg caccgtcatc gccgaggagc aggctcgcac ggtgctgtca 45240
ggcctggggc agccactgac cctggtcgat gccgccagg agccggcctc cgcgcagcag 45300
gtccccgcgc tggggccgga gcgctcgctg ggcggggaca tgctcgccta cgtgcttttc 45360
acctcgggct ccaccggcga gccgaaggc gtctgcattc cccaccgtgc cgtgctgcgg 45420
ctgatccatg agccctccta cgtccagctc tcgccgcggg aggtcatgct gcactacgca 45480
ccgctggagt tcgatgcctc caccttcgag gtctgggggg cgctgctcaa cggggcaagg 45540
ctcgtgctgg tccctccgga acagcagtcg ctcgagagcc tgggacagga gctgagcacc 45600
cagggcgta cctgctgtgt gctcaccgcc gggctcttcc ggctcatggt ggaggagcag 45660
ctcaagagcc tgcgcggcgt gcgccagctg ctgcgccgag gggacgtgct gcccatgccg 45720
caggtacgcc ggctgcgcga ggccctcccc gagtgccagc tcatcaacgg ctacggcccc 45780
accgagagct gcaccttcac ctgctgtcac cgcgtgggca gccccagga gctgggcggc 45840
tccgtgccc tgggcacccc catcgacctg ggctgggtga gcgtgggtgga cgagcggctc 45900
cagcccgtgc cggacggagc cccgggcgag ctgctcgtgg gcggccccgg gctggccttg 45960
gggtacctgc aacaccggga gctcaccgcc gagcgcttca tccccgatcc gctcagccgg 46020
acgccggggg ctgcgctcta tcggacggga gacctggtgc ggcgcaggga ggatgggacg 46080
ctggagttcc tgggcccgtg ggaccaccag ctcaaggtac gcggcttccg catcgagccg 46140
ggcgaggctg aggccgcggt gctcaccac ccggcggtgc agtccgcggt ggtggtgggg 46200
cgcgagggcc ccggcgga ggaagctggc tgctacgccg tgccgcgggt ggagagctcg 46260
gagcagggct cgcagcagga gcagcggctg gtacacgagt gggagtccgt gttcgacggg 46320
cacatgtacc gcgaagcccc cgtgggcgga gagccgacct tcaacatcgt cggctggaag 46380
agcagctaca ccggccaacc cgtggccgtg gaggagatga gggactggct gcgccaccgg 46440
gtggagcgcg tgcgcgggct caggccacga cggatcctcg aagtgggctg tggtagggg 46500
ctgatgctct ttgccctgct gccgcactgc gagcgctacg tgggcacgga tttctccccg 46560
gcggcggttg actacgtgcg tcgctacctc cccccggagc acccgggccc cgtggagctg 46620
ctacaccgca cggcgagcga gtggagcggc gtggcgccg gctccttcga cgcggtgctg 46680
ctcaattccg tgggtgcagta cttcccctcg caggagtacc tgcggcaggt gctggccccg 46740

tgctggagg ccgtggagga cgggggcttc gtcttcgtgg gggatgtgcg gagcttgccg 46800
ctgctggagt ccttccacgc ctcggtggag ctggagcggg ccgccccctc catgccgctg 46860
gaggcgtggc gcgagcgggt gcggcgtgcg gtgctggagg acaacgagct ggtggtggac 46920
ccggcactct tcgtggcgct ggcccatcag catccacggg tgagccacgt ggatatcgag 46980
ctgacgcgcg gcacgcaccc gaatgagatg gcacgcttcc gctacaacgc cgtgctccac 47040
atcgggccgc gcacgccgcc accggcctcc gaggtgccgt ggggtggactg gtccacgcaa 47100
ggactgagcc tcgatgcact gcgggctcgg ctccggcagg gcccgccagg gccactgggg 47160
gtggcaggca ttcccaatgc gcgagtgtg cccgcggtcc gggccgccga ggcgctgggc 47220
agcacaggca gcgcgcggcg ggtggaggag ctgcgccggc gcctgtccca acccggcaca 47280
ggggcccagg atccagaccc cttctggcaa ctggccgagt cgctcggcta cacggcggca 47340
gtgagctggt ctctggggcg gagggacggg gccttcgacg tgctcttcct cccggcgacg 47400
ccagggatgc acccgcggtg gctcggcccc acgcgcgtca accgcacgcc gccccagcc 47460
tcgcgcgctc tgtcctcgga gcctcggcg gccagcctct cactgaggct gggctccgcg 47520
ctgcgggccc acctgcagac ccattacca gacttcatgg tgccctcacg ctctgtggtg 47580
ctgcagtccc tgccgctcac tcccaatggc aagggtggacc gcgcggcctt gcccgctccc 47640
gactcgcgtc ggctggaatc tgccccgctg gtgcccccca gcaacgagct ggagcgggtg 47700
ctggcgcagg tgtggaagga agtgctggga ctggaggaag tgagccggga ggacaacttc 47760
ttcgacgtgg gagggcactc gctgctgctg gcgcaggtgt gcagccggct ggaggcccga 47820
ctgggcccggc ggctcgaact ggtgaccctc ttccgctact cctccattgc tgccctggcc 47880
gagcacctgc agggccccca ggagctggca gccgccgagg cgcaggtgca gcgcatggcc 47940
acggagcgcg ccttgctcca gcaacaggca gcccaacgcc gccggggccg ttccaagcgg 48000
ggagctccca atgacacctg gtggggatga agcgtgcag tcttccatcg ccctggtagg 48060
catggccggc cgcttcccgg gggccccgga cgtggagtcc ttctggcgca atctggtggc 48120
cggagtggag tccatttcct tcttctccga ggaagagctg agacaggcgg gcgtgtccga 48180
gcagatacgt cggcgccccg agtacgtgcc cgccaagggt gtgctggaag acctggagct 48240
gttcgacgcg ggcttcttcg gctactcgcc gcgcgaggcg agccacctgg atccgcagca 48300
acgcctgctg ctggagtgtc cgtgggaggc gctggaggat gcgggccttc gcccgatca 48360
gctcccaggc tgggtgggcg tctacgtggg cgccggggac accagctacc gcttccagct 48420
gctgcgcgga cacggggacc cgctcagcgg ctccaaggat gtggcgggct tctttggcaa 48480
ctaccggac tttctcgcca cgcgcgtggc ctacaagctc aatctgcgcg gccccgcgt 48540
cggcatccac accgcgtgct ccacctcgct cgtctccatc aacatggcgt gcagcgcgt 48600

gcgcggcttc gagtgtgaca tggcgctcgc gggcggcgtc tcgctgcggc tgcccgcgcg 48660
ctccggctac ctctacgagg aggggggagt cgcctcgaag gacgggcact gccgcccctt 48720
cgacgcgcgt gctaccggca ccgtcacggg ggatggcgtg ggcgtggtgg tgctcaagcg 48780
cctggaggac gcgctgaagg ccagggatcc catccacgcg gtgatccgcg gctgggcgct 48840
caacaacgac ggcgcctccc gggctggctt caccgcgccc agcgtggagg gccaatccga 48900
ggtgatcgcc ctggcgcacg ccgcggcggg catcagcgca cgcgacatca cctacgtgga 48960
ggctcacggc accggcactc ccctgggaga cccgatcgaa gtggcgggcg tcacccgcgc 49020
cttcggggcg cataccgctg acaccgcctt ctgtaccctc ggtgcggtca aatccaacat 49080
cggacacctg gatgcggcgg caggagtggc cggggtgatc aagacggtgc aggccctgcg 49140
ccaccggctc atccctccca cgctccactt cgagcggccc aatcccgcgc tccacctgga 49200
gcagagcccc ttcttcgtca acaccagcc gctcccgtgg gagagcccgc gcgggcctcg 49260
gctggcgggg gtgagctcct ttggcattgg cggcaccaac gccacacgc tcttcgaaga 49320
ggctccgccc ccgcccgcga gcggccccac ccgtcccaac caggtgctgc tgctctcggc 49380
ccgaagcacc agcgcgctcg agcacatcg aggacggctc gcggcacacc tgcgccgcca 49440
cccggacctg gaactggcgg acgtggcctt taccctccag gtggggcgag cccgcttccc 49500
ctaccggcgc gccctcacct gtgcgaccct ggccgaggcc atggagcgcc tggaggctcc 49560
cgagccgcg ccacccgagc cgcttgcctc cgagggagag cggccccctc tcgtcatgct 49620
ctttccaggg cagggcactc cgctggctcg cacggcccgt gcgctccatg aatccgagcc 49680
caccttccgc caggcggtgg agcagtgtgc ccggtcctg aggcagacgc tggggctgga 49740
tgtcagggag gtgctcttcc cctccgcgga gcaggaggag caggcacgcc ggctggcggc 49800
gcagacgcgg gtggcccagc ccgcgctctt caccctggag tatgccctgg cgcagacctg 49860
gctgggctgg gggctccaac cgcaggcgct ggcaggccat agcctgggcg agctgggtggc 49920
ggcgtgcctc gcgggtgtct tctccctgga agatgcactt cagctcgtgg cagcccgggg 49980
acagctcatg cagggatgcc caccgggggc catgctggcc gtgccgctgc cggaggccga 50040
gctggcgggc ctgctgggca gcgagctgtg tatcgccgcg gtgaacgggc caccggcctg 50100
cgtggcctcg ggtcctctac cggcggtaga ggcgctcacc gccgcgctgg agagccgggg 50160
cgtgtccagc cggcgcttg agacttccca tgccttccac tctgcctcca tggaggcatg 50220
ccaggggcca ctcaccacgc tgctgcggcg catgcgcctg caggccccctc gcctcccctg 50280
tgtctccggc ctcaccggcc ggtggctcac cggcgaggag gccacggagc ctacgtactg 50340
ggcccgcag ctccgcgagc cggtagcttt ctccgaggcc ctcgagacgc tctggagcct 50400
caaggagcca gtgctgctgg aggtaggccc cggcaccacg ctcaccgcgc ttgcgcggcg 50460
gcaccccacc cgcgccgccc gcacccagga ggtggcgagc cttcccgtgc aaccggacac 50520

28/38

ggccgtgccg tgcattgaag aggccgtggg tgagctgtgg caggccggcc tggagctgga 50580
ctggagcgcc ctgcacgccc cgccacgcca ccgcgcacac cttccaccct accccttcga 50640
gcgccagcgc tactggattg aaccggaggc tgccccgcag ccgcgcgcgc agcagccgac 50700
gcccgcgtcc cttgtgcccc cggagcagcc ctgcgcgcag gccctggaag actggttcta 50760
cgtgcccacc tgggagcagg cacctgccac gagcgggggg ggacagcccc tggcggggccc 50820
ggtgctcgcc ttcattggact cctcgggcct ggccgagcag gtgctggccg cgctgtggcc 50880
cgccgactcc ggcgcgctcc tcacccgcgt cgagccggcc gggcactatg agcaactcag 50940
tgagcacgcc ttccgcctcc gccccgagag cgaggaggac tgggacgcgc tcttccaggc 51000
gctccagtcc cagggccggc tccccgtcg cactctccat gcctgggctc tcaccgcgga 51060
gcccggcccc tgtaccccgg acggggaagc ggtgctggaa cagggcttct tcagcctgct 51120
gcggctggcc cgtgcgctcg gccgccacgc gcctgagcgc cctgtacagc tcgaggtgct 51180
gtccagcttc gtccacgcgg tgggcccgcg ggagccgctc gagccgctca aggccacgct 51240
gctggggggc tgtgcggtac ttcccctgga gtaccctcac gtgcagtgcc gcaccattga 51300
cgtgcggccc gggagcgagc cacgggaagt gctggtgagg agcctggccg ccgagctggc 51360
cgctcccatg ggagagagcc cgggtggcctg gcgcgacggg cagcggtacg tgcgccgcgc 51420
caccggcagc aggcctcgagg cctcgcggcc cctgcgcagc cttcgggagc gaggcgtgta 51480
cctggtagca ggagggtggt gtggaatcgg cctggtgctg gcgcgagcgc tggcccagcg 51540
ggcgctgctc cggctggcgc tactcaccca ctgcaccttc cctccccgag agcagtggga 51600
gcagtggctg gaggaagccc cggcgcaccc ggagccggcg tggcggagcg aggccgaccc 51660
ctcggagcgc cgcaggaccc agcaccgcat ccgctgcctg ctggagctgg agcagctcgg 51720
cgccgaggtg caggtgtaca cggcggacgt ggcgaggag gccgccgtgc gctcggtggt 51780
ggagcaggtg cagccccgtt gggggaagat ccacggcgtg ctgcacgccg ccgccacggt 51840
cgacgacggc gtcattccagc tccgcacgca cgagcagtcc tcgcggggcc tgcgcaccaa 51900
ggtgcggggc agcatggtgc tgcacgaggt gctggcgagc gaggggttg attggttcgc 51960
cctgtgctcc tcgctggcgt cggcgtggg ctcatcggc caggcggact actgcgcggc 52020
caatgccttc caggatgcgt atgcgcacca cctgcgccgg cagggcttca cgggagcgct 52080
ggcgctggac tggggcacgt ggagagatac gggggcggcc atgcggctgg tggcacgcac 52140
ccgacggggg ggccatgaga agccgcccac gccgctcacc caccactct tcgattgcga 52200
acagcgcgag ccgggcggga cgcactggct cggcctcacg ctgaggggtg gagaggactg 52260
ggtggtggac gagcacggc tgcaaggggt gccaaactg cccggggtgg cctacctgga 52320
gctggcgcgg gcggcgtgtg ccagggcgt gggggccgag gcggtggagc tggcgagct 52380

gttgctgctg gagcctctga cggtagcctcg aggcgaatcc aggcagggtcc ggggtggtgct 52440
ccagcccagag gggcagggtc atgccctgcg ggtggagagc cggtcggagg aagcgcgggg 52500
atggaatgag catgcgcggg gccgggtgcg cgccgtgcct cggttggccg agcgcattcca 52560
gcccagagctg ctgcgcgccc cctgcgaaca cgagcagccc gtgcccggag agcccagga 52620
gcaaggccct gtccacgccc gagcacgctg gcatggcctc ttccagtggg ttccgcgcgg 52680
ccctcgccag gccctcgccc agctcgcgct accggagccc ttccacggtg acctcgagcg 52740
cttcgagctc caccggcgcg tcatggacat ggccaccagc ttccgcatte ccggcgaggat 52800
gccctggctc gcctttggct atgagcgcggt gctcattcac ggtccgctgc caccgcaggat 52860
gctcagccac gtgagcctgc ccgaggagtc gcaggctggc gcgcaacaac tccggctaca 52920
ggtgcggtg ctggatctgg agggctggga gcgggtgctg atcgatgggt acctgctgcg 52980
gccgctgaag ccagcgagcg ccagtgtcga gccggcgggc cccaacgtgg aagtcgccgt 53040
ggggacgcgc gggctgctgg agagccttgg cctgcgcgcg tgcaccgcgc ccgcccccg 53100
gccgcgccag gtggagatcg aggtggaggc cgctgggctc aacttcctgg acgtgctggg 53160
cgcgctgggg atgatgccgg cattggaggc ggaggagagc gtactggggc gcgagtgtc 53220
cggacgcatt gccgcggtgg gcgagggcgt cagcgggctg cgcggtgggg acgaggtgct 53280
ggcggtggcc ccaggctgct tccgctccta cgtgctggtg gatgagagcc aggtggtgcg 53340
caggccccgc tcgctggggc tcgccgaggg ggcgggccag atggtgccgt tcgccacggc 53400
gtacttcgcc ctgcacaccg tgggccggtt gcggcgcggc gagcgcattc tcatccacgc 53460
cgcgggccga gggctgggtc tggccgccgt tcagctcgca tcccggaccg gggcgagat 53520
attggccacc gcgggcagtg agcagaagcg cgagtacctg cgctcgctcg gcattgctca 53580
cgtgctggac tcgcgcagca cctccttcgt cagcgaagtg cgcgagcgca ccggcgggcg 53640
tggggtggat gtggtgctca actcgctggc gggagagctg ctctggcgg gcctgtccgt 53700
gctggccccg cacggccgct tcctggagct ggggaagagg gacctgtatg cggaccagca 53760
ggtgggcctc cgaaccctgg ccgagggca gactttcgcc gccattgact tcggtcccca 53820
ccaccggac ttccgagcgg tgctcgagga ggtggccacg caactcacc agggccagct 53880
cgagccattg ccacccgcgc tcttccccgc gcggcagggt gccgaagcct tcagcttcat 53940
ggggcgcgcg ctgcacatcg gtgcgctcgc cgtctccatg cagggggcga cggcattgcc 54000
cggtccatg actcggggct ccaggcccgcc accggtggcc gtacctcctt gggaggaccc 54060
gcggctggcg ggcggcatct cctctgaaga gggtgccgag gccttcctgc gcgcgttgga 54120
ccaggggcca ccgcagctca tcatctcccc ccaggacttc agctcgctgc tgcgcggcct 54180
ggggggcagc cagggcgctg gcgaaaagga gcgcctcgtc accgggcgcg ccgcccgccg 54240
cgagcgcnag gccttgccac cctcctcgct ggagcagctc atcgagcagg tgtggcgcaa 54300

30/38

gcacctgggt	gtggagcgcg	tgcagccac	ggacagcttc	ttccagcttg	gaggagactc	54360
gctgctgggc	atccaagtgg	cggcggatct	ccgcaggcac	ctgggtgtgg	agctgcccac	54420
ggccaccctc	ttcagccacc	ccaccctcgc	cgcgctggcc	gcggctctgc	gggctcgaca	54480
gggtgaggcc	gcggctccca	ctgccccgc	gccagcgctc	gtgccggatc	cagccgcgcg	54540
cttcgagcct	ttccccctca	cggatgtgca	ggaagcctac	tgggtgggtc	gccgctcggc	54600
cttcgagctc	ggcggcgtcg	ccgcccattg	ctacttcgaa	atcgaaagcc	cggggctgga	54660
ggtggagcgc	ttcatccaat	gctggcgcca	gctcctgcag	cgccacgaca	tgctgcgcat	54720
ggtggtgctc	ccggatgggc	ggcagcaggt	gctcgagcag	gtgccggagt	acacgccgga	54780
ggtggtggag	ctgcgggggc	tctccccca	ggaggccgag	tcccgcgggc	tccagctgcg	54840
tgagcgcattg	gcccaccagg	tcctgcgcag	cgaccgctgg	ccgctcttcg	agctggtgct	54900
ctgccggtac	gagggaggcg	tccgcattca	catgagcatg	gatgccctga	tgctggatgc	54960
gtggagttca	gcggtgcttc	ggcaggactt	cgcccagctg	taccacgagc	cgggccggcc	55020
gctggagccg	ctggccatca	ccttcgcgca	ctacgtgctg	gcggagcgcc	ggctgcgcga	55080
gggcgaggcc	catgagcgcg	cccgcgcata	ctggtgggct	cggttgga	cgctgccgcc	55140
accgcccag	ctgcccctgg	tgaaggaacc	ctcgcagctg	gagcacgcgc	ggttcaccca	55200
ccgcgaggct	cggctcgagc	cacaccgctg	ggcccggctc	caggagcggg	cgcgcgcca	55260
cggcctcacc	ccctcgccg	cctgcatggc	cgccttcgcc	gaggtgcttg	cccgtggag	55320
ccgtcaccgg	cgcttcaccc	tcaacctcac	cctcttcag	cgcttgcccc	tgacccgca	55380
ggtggacgag	ctggtggg	acttcacctc	cctggtcctg	ctggaggtag	aggcacacgc	55440
ggcgagcacc	ttcgccgagc	gtgcctcccg	gctccaggca	cagctatggc	gggacctgga	55500
gcacggcagc	gtgagcgccg	tgcagctcat	ccgcgagctc	gtccgcaccg	gccgccgctc	55560
cccgggcgcc	atcatgccc	tcgtcttcac	cagcatcctc	agcctggatg	cgcggcgcg	55620
ccccagggc	agcctctcct	tcttcgagg	agaactgggtg	tacagcatca	gccagactcc	55680
ccagggtgtg	ctggaccacg	gagtccacga	ggaggagg	gcgctcgctc	tggcgtggga	55740
ctcggtggag	gcgctcttc	ctccgggcat	ggtggacgac	atgttcacg	cctaccagcg	55800
gctgctggg	gcactcgccg	aggaggagca	agcgtgggag	ggcgagctgc	cggagctgct	55860
gcctcctgcc	cagcgtgagt	tgctcgcg	ctacaacgcc	accagggcg	cgcggcccag	55920
cggacggctg	gaggagggt	tcttcaccca	ggcgcggctc	caccccgagc	tccccgcgct	55980
gctcgacccg	gagcgcaccc	tgagctacgg	cgagctggca	aggcgagccc	aggcgctggc	56040
cgcgcgccta	cgcgagctgg	aggtgcagcc	tcaggagttg	gtggccattg	ccatgcacaa	56100
gggctgggag	caggccaccg	ccgtgctcgg	cgtgctgcag	gcggccgcg	cctacctgcc	56160

gttgatcca gagcagccgc cgctccggct ccaccagctc ctggaggagg ggcccggccg 56220
cgtggtgctc acccagtcct ccttgctgca caccgtgccc tggccgcccg gtgtgcaggt 56280
gatcgccgtg gacgaactcg agcctgcaac cgaggccccc ccgctaccgc ctgcggttac 56340
accggagcac ctgcctacg tcattctatac ctccggctcc accggcaaac ccaagggtgt 56400
ggcattgaa caccggggccg cgctcaacac cgtggtggac ctcaacaccc gctttggcgt 56460
cggcccggag gatcggtgctc tgggcctgtc cgcgctcacc ttcgacctgt cgggtgtacga 56520
cgtgctgggg ctgctcggcg cgggcggagc gctggtgctg cccgcccggg aggcggagaa 56580
ggatcccgc cactggtggg agcggctggt ggctgggcgg gtcacggtgt ggaactccac 56640
cccggcgtg atgctgctgc tgggtggagta cgccgagcag cgccggctga agctgcccgc 56700
ggcgtgctgc ctggtcatgc tcagcggcga ctggatcccc gtggcgctac cggatcgcat 56760
ccgcgcactc ggaggggacg tgcaggtggt gagcctggga ggcccaccg aggttccat 56820
ctggtccatt gcctacccca ttggccaggt ggcaccgcag tggaagagca ttccctatgg 56880
catgccgctg gccaatcagc gattccacgt gctggacggg cggctggagg cacggccctg 56940
gtgggtgccg ggcgagctgt acattggcgg cgaggggctg gcacgcgagt actggaggga 57000
tgagccgctc accgcaacgc gcttcacccg tcacccgcgc acaggcgagc ggctgtaccg 57060
caccggggga caagggcgca tgctgcccga agggagcatc gagttcctcg gccgggagga 57120
tctgcaggtc aaggtgcagg gcttcgcgct ggagctgggg gagattgaag cggcgctggc 57180
ccagcaccgc gcgctatcgg cgagcgtggt ggtggcacga ggagagcctc ggggagtgcg 57240
gcggctggtg gcctacgcgg ttccctcgctc gggtcagacg cccgcggcag gcgagctgcg 57300
ccgctacctg gcggagcggc taccgcgta catggtgccc tccgccttcg tcctcctgga 57360
gtccctgccc cgcagccgca acgggaagat cgcccgagat cagctccccg agccgcagca 57420
gacacaggga ctggctgcac aggcgcggc ggccgacccg ctcgaggagc ggctggcagc 57480
cctggtgaaa gaggcgctga ggctcgagcg cgtggagccc caggacagcc tgctggacct 57540
gggggcggac tcggtggcgc tcattccgct catcaaccgc ctggaggcag agctgcagtt 57600
ccgcccgcga ctggccgaca ttacgagaa cccaccgtg caggggctcg ccacgctcca 57660
ccaagagaag acaaagagcc agggagaggg aggcgctccg cggtcacgg cgcgccgctc 57720
cacgcttctg cccgcccagg agtggggcg cttcaaggcc aaccgcccgg gcctgcgccg 57780
cttccccgat ggaacgcctg aggtggcact gccaggcagc ggcttggcgc cggccccga 57840
ggagctgaca gctctcagc gccgcccagc tgtgcgcacc tactccctcg agcccgtgag 57900
ccatgagcag ttgggacggc tgctggcccc cttgcgcgag tgggaggtgc agggctcgcg 57960
gcgttacctc tacgctccg cgggcgggct ctaccgggtg cagctctacc tccacctcaa 58020
gcccggccgg gcgcgaggc tcgagcccgg cacctggtat tacgatcca gcacccatcg 58080

32/38

gctggtgctc ctgtccgccg gagccgggct ggatcgccgc atccatgatc cgcaccagaa 58140
ccaggccatc ttcgactcgg ccgcgttctc cctcttctc atcgcccgca tgggggccgt 58200
ggagccggta tacgccgaac atgccctgca cttcgccacg ctggaggcag ggctgatgac 58260
ccagcttctc gatctgggcg cggctcccag tgggctgggg ctgtgccata tcggcgacct 58320
ggacttcgcc caggcgccgg ggctcttcca cttggaggag gagcacgtgc tgctgcacag 58380
cctggtgggc ggagtactcc ccactcgggg gcaggaggcg gcctcgggtc ccgctgaggg 58440
aggaacggag gcgcggcaac tggcgagct cctgcaacag gtgaagacgc tcacgcccga 58500
ggccgcgcga gcgctgctcg aggcccgccg cggcagcaag gggagaccac atgagtaagc 58560
tgacagagga gctggagagc ctggcgccgg agcagcggga gctactcgcg gccctgatga 58620
aagagcaggg cctggacgag ggcgcattgc tgatgcccggt ggagcgcaag cccgaggggc 58680
tgccgctctc gtccgcccag cagcgcatgt ggttcctgca acagctcagg cctacgagcc 58740
ccttctacaa cgtgcacgcg gcgctccggc tcacgggaaa gctgaagggtg gagtgcctgg 58800
tgacagacct gaatgagttc gtccggcgtc acgagccgct gcgcacgggtg tccccctccg 58860
ccggaggtca acccctgcag cgcatacctc cgcccgacc ggctgcgctc gagcagcgcg 58920
acctgagcgg cgtgcccgcc caggagcgag aggccgaggt gtaccgcgcg gtggagcacg 58980
cagcgctcgc ctcttttgat ctggagcgag agccgccgtg ccgcttctc ctggtgcgag 59040
tggagccgga ggaacacgtg ctggtgttcg ccacgcacca catcgacgcg gatggctggt 59100
cgctcggcgt cttcgtgcag gagctgtgcg cgctgtacac ggctgctgta cacgcggagc 59160
caccggcgct gccaccgctg cggctgcagt acgcggactt cgccgcgtgg gagcggagcc 59220
ggctcaaggg tgggcgcgag cgcgagctgc tcgagtactg gcaggagcag ctggcagggc 59280
tgccggacct gagcacctg ccgccagagc ggccaggcc cccgctgtcc aagcaggagg 59340
gcgccacgtt cgaattcgcc ctgccaccac accaggtgca ggccctgcgc tcacttgctc 59400
aagcgcgggc caccctcgctc ttctccgtgc tgctcgccgc cttccaatgg ctgctggccc 59460
ggtgcgcccg ccaggacgac gtggccctgg gcatgcccat cgccaaccgc aaccgcaagg 59520
agctggaagg cctcatcgcc tgcttcgcca gtaccctggt gctgcgcgcg aagccctccg 59580
cctcgaccgc cttcacctca tggtcgcgcc aggtgtccga gcagctccac ggcgccctcg 59640
agcaccagga ggtgcctttc gaacgattgg tggaggtgct ccagccccgg ccgcgcatgg 59700
atcggcaccc gctcttccag atcttcctcg ccatgcagca gcaccgctg cggcgcgccg 59760
agctgccagg cctgctgctc tcggagttcc cccttcgag ccgctgggcc cgcttcgatc 59820
tggagttcca cctctgggag tctgcccagg gcgtcgaggg caccctcatc tacgacgtgg 59880
acctgtatgg ccaggagggc gtagcgcaac tgggtcgggc ctacgtgagc ctgctcgaag 59940

ccgtggcggc ggacccgacg cggacactgg gggagctggc gggtagagacg ctggccccgc 60000
ccgtacggaa gcaggttctc gccctgtcgg agtgtccgcc cctcccacgc ccgagcacag 60060
cccctcggac actggccgag gcacttctac agaccgccga gcgcttcccc acggccacgg 60120
tcagcttcgt ccaggcacag ggctcttgca cggcctggac tctgccagag ctggtggagc 60180
gggcccggcg cctgcaggcc ggcttgccgc agtggggatt gcggccgggc gacagtctgg 60240
tgctggtgct ggggagggag gaagagacgg tagaggcact ctgggcctgc gtgctggcgg 60300
gagtggcgcc cctggtgcta ccggcgccgc ccgcgagagc ggaggcgagt cccgcgctct 60360
cacgtctgcg gcatgcccgg cagctgctgg gtgggccccg ggtcctcacc cggcaggaga 60420
tgctgcccga tcttgccagg cagcttcagg tctcgccac cgccgacatc ctgggggccc 60480
tgaggagct gagagccacg ggtggcgagg cggccctgcc ccaggggcg atggatgacg 60540
tggccctgct caacctcacc tcgggcacca ccggcaaggc caagtgcgc atgctcacgc 60600
accgcaacct gctggtgcgc ctggaggcca ccaacgtcgt ctacgagtcc cagccgctcg 60660
agcgcgccct cgtctggttg cagctgcaca acattggcg cctgtccgag taccacctcc 60720
gcccgtcttg cgccggcatg cacaccttc acgctccac cgaggaggtg ctgcccagc 60780
ccctgcgatg gctggagtgg ctcgagcgg atggcattgc acagacgtgg gcgcccagct 60840
tcgcctacag ccacctgctg gagcggtgc gcaaggtgga ggaccggcg tggagcctgg 60900
gaggcgctgc cgtgctgctc agcgccggag agcagatctc cgcgcccatg gtggaggagc 60960
tgatgcggag gttggctccc tcgggagtgc gggaggatgc ctctgctcgt gcctggggaa 61020
tgaccgagac ggcttcgggg gtgacctatg cgcgggcgcc cggaacgccg ccccgcatgc 61080
acacctgga gcgcgcgagc ttgagcggtc cgcttcggca cgcggcaccc gcttcgccc 61140
ccgcgctgcg gctgatggat gtaggagccc ccattgccgg caccgccctg cgggtggtgg 61200
acgcgagcgg cgagctgctc tccgaggagt gcgtggggcg cattcaagtg cgcggcgaga 61260
tgatctctcc cggtactac ggggacccga aggcctcggc ggcgctgctc accgccgacg 61320
gctggctgga gacaggcgat ctgggcttcc tctccgagg agcgtgacc attaccggcc 61380
gcgccaagga tctggtcatc atccacggca ccaacttctc ctgctacgag atcgagtccg 61440
cgggtggagca ggtggaggga gtggcgccct cctcggcggc ggcggcggcg gtgcggatgc 61500
tcgaggggag cagggaggag ctggcggctc tcttcgtgcc cacggaggga ctgctccgc 61560
aaccctctgc ctactcctg tcgcgcatcc gccagcagg gctcgagcag gtgggggtgc 61620
ggatcgatca tctcatccc ctcgagcctc accagctccc acgcacggag ggaggcaagc 61680
tgcgccgctc ggagctgagg gcgcgcttcg aagcaggatga gctgaggcg ccgcaaccgc 61740
ctccggtccc gagcccctcg cgcccgctgg agcagctcat cgcctccgtc tgggcggagg 61800
tgctcgagca ccaggacatc gcgcccgagg ccagcttctt cgacctgggc ggcaactcca 61860

34/38

tcctcctggt gcgcgtggag cgggcactgc gcgcgcggct cggcctggag ctgacgctca 61920
tggacctgtt cgcctacccc acggtgcact cgctggcgga ctacctggag cctcgagccg 61980
cccagctgcc agcacaggcc tccccctcca cgcaggccga gcgccggcgg gggatgcgcg 62040
gcgcggcgct ggagcagaga cgcactcgcc gccaggcaca acgggacgag gactagacac 62100
gcatgagcca cgacacgcag ctcgagggag ccgtagcggc agtggccatg gcggggcgct 62160
tccccaggc cccaccctg gaggactact ggcgcatcat ccacgagggc gtggacgcct 62220
ttaccagct cacggacgaa cagctcctgg cctccggagt ggggccctcg ctctgcgcc 62280
aaccgggta cgtgcggcgc gccgccctgc tggacaggat ggaccagttc gacgccgcgt 62340
tctttggctt ctccccgca gaggcggagg tgctcgacc acagcaccgg ctcttcctgg 62400
agtgcgcgca cgaggcgctg gagcgggccc gccatggctc cgagcgcgca agggggcgcg 62460
tgggggtctt cgccagcgcc agcctcaaca gctactacct gcattccctg cacgggaacg 62520
cccggctgag ggaggtgctg ggagacttcc agctcgccat cgccaacgac aaggacttcc 62580
tgccacccg cgtctcgta aagctgggcc tgcgcgggcc cagcgtggcg gtgcagacgg 62640
catgctccag ctcgctggtg gccgtccacc tcgctgcca gagcctgctc aatggtgagt 62700
gcgccttggc gctcgcgggc ggctcctctc tctccgtgcc ccaggcccag ggctacctct 62760
accaggaggg aggcattgcc tcaccggacg ggtactgccg ccccttcgac gcggcgcgcg 62820
cgggcaccaa ccgaggcaac ggctggggg tgggtgctgct caagctgctc gaggaggcgc 62880
tggcggacgg agacaccatc cagccgtca tccgcggtc ggcggtgaac aacgacgggg 62940
cacacaagat cggctacacc gcgccaagcg tggaggggca ggcctcggtc atctccgagg 63000
cgctggaggc ggcggccgtc tccgcgga gcatcggtc cgtggaggca cacggcaccg 63060
ccacggccct gggagatccg atcgaggtcc aggcctcac gcgggccttc cgcaagcaca 63120
cggagcgccg cggctactgc gctctcggt cggcgaagg caacatcgga cacctggatg 63180
cggcagcggg catcgccagc ctgatcaagg ccgtgctggt gctggagcgc cgccagctcc 63240
cgccatgccc gcacttcacc tcccccaacc cagcgcgcga cttcgagcgc agccccctct 63300
acgtgagcgg gcgggggag ccgtgggagc cggcggacgg cccgcggcgc gcgggcgtca 63360
gctccttcgg tattggtggc accaacgtgc atgtggtgct ggaggaggcg ccgccccggc 63420
cggcgccgcg gcgggcagcg cggccctacc agctcctgcc gctgtcagca cgcacggagc 63480
gcgcccggga ggaggtcgag gcccggtgc gcgagcacct gcgccagcag cccagggagc 63540
cgctggagga ggtggccac acgtccagg taggcgcgg ccacatggcc tggcgcacag 63600
cgctggtgag cagcagccct gccagggcg tagagctgct ggagaggcgc cgacccgagg 63660
cgctgctgcg tggccagagt gcggcccagg ccgctccgt ggccttcctt ctaccgggccc 63720

agggctcaca gtacgtgggc atgggagcgg cgctccatga gtccgaggcg cccttccgtg 63780
aacaggtgga tctctgtgcc gggaaactcc agggcggtgct cgggttggac gtgcgccgcc 63840
tgctgtaccc tggcccgcgg gagcgggaat gggcccagga gcgcctgcgg gagacgcggg 63900
tatgccagcc ggtgctcttc accgtggagt acgcgctggc ccgactgctg gaggcctggg 63960
ggatacgcgc caccgcgctg ttggggccaca gcctgggcga atatgtggcc gcgtgcctcg 64020
cgggcgtggt ctccctggaa gaggcgctgg aggtgggtggc cgcacgtgga cagctcatgg 64080
gcagcatgcc cccggggggc atgctcgccg tggggctccc agcccaggaa gtggaaccgc 64140
tcctaccgcg ggcaactggc ctggccgcgc acaacagccc ccaggcctgt gtggtggcgg 64200
gccccaccga gcccctggcc gagctgcgcg cagcactcga gcagcggggc atcgcgtgca 64260
cgccctctggc tgtctctcac gccttccact cgccgatgat ggagccggcg gtggcccat 64320
tcgtgccccg gctgcgccgg atgcatctgc ggccctccag cctgcccttc atctccaatg 64380
tgactggcac ctggatagag gccgaggagg cgacctcacc ggagtactgg ggccggcacc 64440
tgctgcagcc ggtgcgcttt gcccaggggc tggagcgggt gtgcgagggt gtgcagccgc 64500
ggctgctggt ggaagtgggg ccggggccaca cactggggcg gctggcggcg cgccaaagct 64560
ccggccccgt gcaggtggtg tccacgctgg gctccagccg ggaggaaagc tcggaggtgg 64620
agcggctgct gacggcaatc gggcgactgt ggggtggagg agccgagggt gactgggctg 64680
ggctccaccg cggcgaacgc cggaggcgcg cggtgctacc cacctaccct tttgagcacc 64740
agcgtactg ggtggagaca gccccccagc ctccccctcc cgagcggcct gggacacctg 64800
ccgagagccc tgtcaccagc agcttctacg tgccaggttg gagccgcgcg gctctgccct 64860
cggcggcccc ctcgagacgg gtggggccggc tgctgttggt ggcgagggcc catgggtgga 64920
gccaggggct cgcgagcgg ctccggggccc ggggacactc cgtcacctc gtcgagccag 64980
gtgagcggct cgagcgcctg accccggagc actggcgcct gccccgggc cgccgcgagg 65040
acttccaacg cctgctggag gactccggag agctgccctc gaggggtgctc cacctctggc 65100
tgctggggag cagccagagc ccccaggagc gcggcttcta caccctgctg gcgctggccc 65160
aggcactggg agcccacggc acccgggccc cggtggagct cacggtggtg accgatcagc 65220
tcacgcggg ggcagagggc gagccgatac agccgctcaa ggcgctcctg cagggccctg 65280
cttccgtgct tcccaggaa ctccccgggt gcacgtgccg cctcgtggac atggccctgc 65340
cgccaggcgg cgtggcggag gagctgctgg agcggctggt ggcgagggtg gagtccaccg 65400
cqtccaaacg ctcggtggcg taccgcggcg cggtcgcgtg ggtgcgcgag ttcgtecccg 65460
ttcgctggcc cccgcccag ccttcgcagc tcccgttcg gcagagaggg gtgtatttga 65520
ttctcggcgg cctggggggc gtagggctcg cgctggcaga gcacctggcg cggcgcgtgt 65580
cggcgcggct ggtgctcacg ggccgctcg ccacgccacc gcgcgagagc tggagcgcag 65640

36/38

ggctgggcac tcccacgcgg ctgcggctct cacaggagct cgagtggctg cgcggtgccg 65700
ccgagcagat cgagcaccag cggccctgc actccctggc cgacagtccc ggcctggaag 65760
agtcgctgcg cctgctgtgc gccagctacc tctaccaact tctcttccct cttcaggcgc 65820
ctctgagact tggagagccg agggcgatgc aggcattgcg tgagcgctg gggctgcagc 65880
ccggcttcga gcgcctcttc tccttcatga tgggcacct ggagcaggcg aagctgattc 65940
gagtgggaga gggcatgctc gagggtcagg tggagccagc ccacgtgccc acacctcgag 66000
cactgcacga gcgactgctc gaaggctacc cggaggtgtc aggcctgctg gagctgcttg 66060
agcattgtgt gcgccactac cctgaggtgc tccgaggag cctgccggcc ctgagtgtgc 66120
tctacccctc gggccgctcc gaggaggact ccgcacgaag cggggtcgag tggagcagca 66180
tggggcagtg cgtggcgctg ctgcggcatt tctctgccc ccatgccgag cgcacgcagg 66240
ggcgtaacct gcgcatcctc gaggtgggcg gcggcagtg cgtggtgctc caggccctgc 66300
tgccgctgct gcggcaccac ccggtggagt accacttcac ggatattggc cctccttcg 66360
tccgcgccat ggaggggagt gcccgccgcc agggcctcac ctctctcac acctcggtgc 66420
tggacatctc ccgccctccc ccagagcagg gccatccgcc tggctcgat gatctggtga 66480
ttgccctcaa cgtggtgcac gccacccgc gcgttctca gtccctggcc cacctcgaga 66540
gcctgctcgt ccaggcggc cacctgtgcc tgggtggagc ggtgaagcag cagccctggg 66600
tggacatgat ctggggtttg gccgaggggt ggtggagcta cgaggacgag ctgcgcaccc 66660
gctctccctt gctggaggta ggcgactggg agcacgccct gcgcgacgtg ggctttgccg 66720
aggtggaggt gctgccagcg gccgtggagc agcgtcacg ctgggacaac gtgctcctca 66780
ttgccagcg gcccgcgag agcgggcttc agccagtggg cgggcgcgcg gccatgcagg 66840
agcggatccg ccggtgcg gccattgaag aagcgggagg cgaggtgctg cccctggtgg 66900
cggacgtcac cgatcgggag cgcattggcg aggtgctgc cgaggtgaag gcccgccatg 66960
gggcgctcca tggggtgatc cacgccgcgc tgggtgctgga ggatgggctg atgcagctca 67020
agacgcgca gtccggccgg cgcgtgctg cctccaaggt cgagggcacc ctggtgctgg 67080
acgaactgct gcgcgacgag cactcgaact tcttcgtcct gtgctcctcg ctgagcgcac 67140
tgctcggcgc gctggggcag gcggactatg ccgccgccag tgccctcctc gacgcgtatg 67200
cgcactcaca gcgagggcgc acggaccgcc gcaccatctc catggactgg gatcgctggc 67260
tggaggtggg gcccgccatg gcctggggc tggggctggc ggcgggagca ctggggctgc 67320
agcgcacggc acccgagag tacacggtgc gatggcaggc cgagcgctgc tgggtggctgg 67380
acgagcaccg cctggaagga cgggccacgc tgccgggggt ggcgtacctg gagctggtgc 67440
gagccgccct ggtccaggaa ctgggcgagg ccccggtgga gctggagcaa ctggtgctgc 67500

tctctctgct ggaagcgccc gcgggggagg aggtggaggt gcgattccac ctgcgaccgg 67560
aggatgaggg ctatgcgctg gagatccgca gccgggctgg cggcctcgcg aatgggggct 67620
ggcgcagcca tgcgatgggc cgggtgagga tgctgcctcg cgggagcgct cgaccacccc 67680
acccgctacg ggagctggag gaacggctcg gcttgacca ggccgaggg gagcatgagc 67740
cagctcaggg ccccgcgag ctggcacctg ccctgggccc gcgctggtcc tcgctgagct 67800
ggcaccgcgc gtggaagga gaagagggcc tggccttgat cgagctcccc gaggagctgg 67860
ccgaggatct gcggcaatgg cccctgcacc ctgccctgct ggacgcggcg accgggtttg 67920
caccgctgcc ccccggtgcc tggctgccgc tgtcctatgg caacacccgc atccacggcc 67980
cgctgccgcg gcagctgtac agccacatcc gccggctgga gcccgccagt gccaggccg 68040
gagtgggtgcg gctggaggtg cgccctcatg acggcgaagg gcgggagctg ctgtgcgtgg 68100
aggagttcgt gctccgccgg gtggaggtgg actcgttggc tcggccacag cccgctggca 68160
aggctccgcg ccaggcagtg cttccacgac caggcgcgct cgactcgctg cgactgcaac 68220
ccctggagcg ccttcgccg caggagggag aggtggaggt gcaggtgctc gccgcgggccc 68280
tcaatttcaa ggacgcgctg ctggcgctgg gggcgttgcc agtggagctg gccaacggcg 68340
ccccactggc gctgggagtc gagtgcgcgg gcatcatctc cgcggtaggt ccaggagtga 68400
gaggcctgcg agtgggagc gcggtggtgg ccgcggcagc aggagccttc gcctcgcatg 68460
tgcgagtccc ccaggagcag gtcttcccca agccagcggg cctgagcttc gagcaggcgg 68520
cgatggtgcc cgtcactctg ttaccgcct ggtatgccct ggaggagctg gcgcggctgc 68580
gcgcgggaga gcacgtcctc atccacgccg cggccacggg ggtgggactg gcgggggtga 68640
agttggcgct gcggcggggc gccacggtgt acgccaccgc gggcagcgag ccgaagcgcg 68700
agctgctgcg ctccctgggg gtaacgctgg ccatggactc gcgcgcgccc gggtttgacg 68760
accagcttct gcaacacacc caggggagcag gcgtggacgt ggtgctcaac tcgctgagcg 68820
gcgagttcct ctcccgagc ctgggcgtgc tcgctcgcca cgggcgcttc gtcgagctgg 68880
gtgtaaggga catgctctct ggcggcacgc tgccgctggc tcccttcgag cgcgggctca 68940
ccttcctcgc ggcgagata gacccgggga tgaagggcta ccgggagctg atgggagagg 69000
ccctgcggca gatcgagcgc ggagagctgg agccgctgcc ctacacggca tggccactgg 69060
agcgcggtggc cgaggcgctc cagctcgtct ccaagggcag gcacgtaggc aaggtggtgc 69120
tcaccccgga agagcccttg tcacggccgc cagcgctccc gagctcctca aggccagtgc 69180
aaccgcgggc aggcgtggcg attgtcgggc tgcgctccgc cgagggtgc gaggccttcg 69240
agcgcatcct cgcggcgggc ctgccgcagg tggccgtctc caccgctgag ctgcgcgcgc 69300
gcatggcgga gatcgagcgc ctgcgcgtct cctcctggga gccggcgctt ccctcggtgc 69360
cccgcggtc ccccgccagg accgaacgct ccaagcccta tgttgctcct cgcaccgagc 69420

38/38

gcgagcgcgc cctggccgcg ctctggcagg acctgctcgg cgtctccgag gtgggggttg 69480
atgatgactt cttcgagctg aaaggggact cgctgctcgc catccagctc atggggcggt 69540
tgcgcaagga gctggaggtg gaattgccac tctcctcctt cctcgcccgc cctacgctgc 69600
gcacactgct ctccgccatg ccagccccgg caacggacga gcctgcctcg gccaccgggg 69660
caatcccagc cgccgctccc cagccgtcac cggtgtctc cagcgaacct gctcggaagt 69720
ggcgccacct ggtgcccata caacctcagg gcgagagccc tcccttcttc tgggctgcgc 69780
cgctgatggg caccgtgttt ccctacttca cgctggcccg ctgcctggcg cccaccacc 69840
ccttctatgc tctccagccg ccaggggtgc aggaggggca ggagcccatg gacaaggtgg 69900
aggaactggc ggccctgtac gtgcgcgagc tgagagagct gcagccgcga gggccctacc 69960
gcctgggagg ctggtccttc ggctgcgtgg tggcttatga ggttggccta cagctcgagg 70020
cggtggaga gcaggtggcg ctgctgagcc tgctggactt ccctgcccct tcggggccagc 70080
gagccagtct ctgggcgacg gcccgttct tcagcggctc ggtgctgcgc ggcctggcgc 70140
cctatacgt cgactatctc tacctggcgg ctgcgcctcc tctgaatcca gacgtgcct 70200
cgcgctgag cagctccctg acccgaggat gggagcgcct gcggaagga gggctggtgc 70260
aggaactgct ggaccatgcc gccatggccc gggtggtgcc tgaggactca cggctgctgc 70320
tgctgcgaga gcctggcatt gcccctatgt tgaggctgac ccgggctcac cagcgagcca 70380
tgctctccta tcggccccga ggtcgcctcc gccagcgcct cgtcctcttc cgcacctccg 70440
agcaggcctg gagcttcgcg cgagacctgg gctggggagc cctcagcgc tcgggcgtgg 70500
atgtgcgcga ggcgcggga gaccatatga cgctgctgcg gccacctcac gtggagcggg 70560
tggcagagca gctccggggc ctgctcagct ccgcgcctct cccgactcgc cgctgagttc 70620
gggctgcgcg ccagaggctg cctcatgcat gccctttggc agccagcccg agctgtccac 70680
ccactccagc gcctcgtca gatcgtgctt ttcgacgac tgatccaagg tgaccacctg 70740
cgcttcacg ctgttgccca gatgagagg caccgcgga tc 70782

SEQUENCE LISTING

<110> Gesellschaft fuer Biotechnologische Forschung mbH

<120> Tubulysin-Biosynthesegene

<130> 13289

<150> DE 102 41 152.7

<151> 2002-09-05

<160> 46

<170> PatentIn version 3.1

<210> 1

<211> 82868

<212> DNA

<213> Angiococcus disciformis

<220>

<221> CDS

<222> (5719)..(7164)

<223> ORF14

<220>

<221> CDS

<222> (12841)..(13881)

<223> ORF11

<220>

<221> CDS

<222> (14942) .. (15586)

<223> ORF9

<220>

<221> CDS

<222> (15847) .. (16983)

<223> ORF8

<220>

<221> CDS

<222> (22366) .. (23532)

<223> ORF6

<220>

<221> CDS

<222> (24591) .. (26513)

<223> ORF5

<220>

<221> CDS

<222> (26597) .. (27517)

<223> ORF4

<220>

<221> CDS

<222> (29858) .. (30400)

<223> ORF3

<220>

<221> CDS

<222> (31220) .. (32392)

<223> tubA

<220>

<221> CDS

<222> (35522) .. (40147)

<223> tubB

<220>

<221> CDS

<222> (48011) .. (58558)

<223> tubD

<220>

<221> CDS

<222> (62103) .. (70616)

<223> tubF

<220>

<221> CDS

<222> (75488) .. (76645)

<223> ORF20

<220>

<221> CDS

<222> (76787) .. (77545)

<223> tubG

<220>

<221> CDS

<222> (77769) .. (78695)

<223> ORF21

<220>

<221> CDS

<222> (79138)..(80019)

<223> ORF22

<400> 1

gcggccgcct tcttcaccgc gggggcctcg gcctcctcgg tggtcggggc gggcgccagc	60
ttgtcgccgt gcaccttgcc gggagcggtc gccttcttct tcgaggcgga gcgcacgggc	120
tcgccctgcg tcacctcgcg cgtgggcagc atgtccccgt gggccttcac cgggcccggac	180
tcacctcct ccgcctcggc gccctcggcc ggcgtctcgc ggggtgtcgcc accctccttg	240
gcccggttga actcgcgcat cgcgtcgtcc accagcccca tggtcattga cgccacgccg	300
aggtcatggt ggtcggatgg cgacaggccc tccttcgtgc cctcgcgag cttcgcgagc	360
gcgtcctgca cctgcgtgtc cccggcctcg gacgcggtgc cggctctcgcc cagttcaccc	420
ttcagctcct ggcccaggtc cagcccgccc ttctcttctt gattcggggc gaccttcacc	480
tgggcccggg cggaggggga ggcacctct tccgcggtgg ggggtgtgct ctcgggagtg	540
gtgctctcgg acggcgccgt ctctggcatg gcgggctccg gggcaatccg ctgcaggtag	600
tcctgaagct tggccttgcg ctcttcacgc tcctcgacgc gggcgcggtc cttctccacc	660
acgtccgggg gcgccttggc cacgaagtgt gggttctcca gcttgcgag cagcccggcc	720
atctcctgct cggcgcgggc aatctccttc ttcaaccgct cgcgctccgc gtccaggtag	780
acgaggcccc ccagcgggac gtagatctcc aggttggagc ccacgaaggc ggcggcctgc	840
ggcggttgg cgcccgggc tcccaccgtc acctcgga ggcgggccag cggcatgagg	900
tagccgcgcc agcgtccag cagctcgcgc gtgcgcacgt ccgcgctctg caccaccgcc	960
ttcaccttgg tggcggggga caggttgctc tcgccgcgga tgggtgcgag gccctcgatg	1020
gcggcgatga ccggcgccat ctgccttcc gccgcctcgt ccacgtgcgc cgcgtccggc	1080
tccgggtacg gggcaatcat gatgctgtcc gtcggccggg ccacggcag cttctgccag	1140
atctcctcgg tgatgaacgg catgaacggg tgcagcagcc gcaggatgag gtccaggcag	1200
tacacgagca cccggcgcggt ggagtccttc gccgcccgt cctcgccgta cagcgagccc	1260
ttcgccagct cgatgtacca gtcgcagaac tccgccaga ggaactggta gagcgtggag	1320
gcggcctcgc cgaagctgta cgactccagc agcgcgcgcg cctcgtgggt ggcgcgctgg	1380
aggcgggaga gaatccagcg gtccgccagc gtcagtgtgc tccggtccag cgggcccggc	1440
tccatccgga agtcgcccac gttcatcagg gcgaagcggc tggcgttcca cagcttggtg	1500

ccgaaggcct tgtagccctc gagccggtcc atcgacagct tgatgtcgcg gccctgctgg 1560
gtgagcgagg cgagcgtgaa gcgcagcgcg tccgcgccga acggcgggcat gccctgcggg 1620
aaccggttct tcagcgtcgg ctggagcttg tcggccgggg cgccgaggac gatgtccagg 1680
gggtcgatga cgttccccctt cgtcttggac atcttctggc ccttctcgtc gcgcaccatc 1740
gcgtgcaggt acacggtgcg gaagggcaca tcccccatga agtggatgcc catcatcatc 1800
atccggggcg cccagaagaa gatgatgtcg tggcccgtct ccatgacgga cgtcgggtag 1860
aaggtcttca gctccggcgt ctcacgcggc cagcccagcg tggagaacgg ccacaggccg 1920
gacgagaacc aggtgtccag cacgtccggg tcctggatga aggacgcgcc gccgcacttc 1980
gggcaggact tcggctgctc gcgcgcgacg atgggctccg cgcgcgcgaa gtccaccccg 2040
cccaccttca ccgtgggggc gtccagcggc aggtccgtgt cgtcgccctg ccgcgggctg 2100
cacgaggtgc agtagtacgc gggaatctgg tggcccacc acagctggcg gctgacgcac 2160
cagtcgtgga tgttgcgcat ccagtggaag aacgtgttcg tccaggactc ggggacgaac 2220
ttcgtgcggc cctgctccac cgcctcgatg gccggcttcg ccagcggttc aatcttgatg 2280
aaccactgcg gagacaggcg tggctccacc accgtggcgc agcgtggca ggtgcccacg 2340
ttcagcatgt ggggctcttc cttctccagc agcccctgct ccgtgaggtc cgcagcacc 2400
tgcttgcgcg cctcgaagcg atccatgccg gcgtacttgc cggctctcctt cgtcatccgg 2460
gccgcttcgt ccaggatggt gagcatgggc agcttgtgcc gcaggcccgt ctggtagtcg 2520
ttgaagtcgt gcgcggcgt caccctcacc acgccggtgc cgaacttcgg gtccaccagc 2580
tccgcgtccg cgatgatggg aatctcgcgg tcggtcagcg gcagcttcac gctcttgccg 2640
gccagcccca ggtagcgctc gtccctccgg tggatggcca ccgcggtgtc gcccagcatc 2700
gtctccgggc gtgtggtggc cacggtgagc gtgcggtcgc tgtccttgac ggggtagcgg 2760
atgtgccaga tggagccctt cttcgactcc tcgtgctcca cctccaagtc gctgagggcg 2820
gtgcggcacg aggggcacca gttgatgagc ttctggggcc ggtacatcag gccctcttcg 2880
tacaggcgca cgaagacctc gcgcacggcg gcggaggact gctcgtccat ggtgaagcgc 2940
tcgcggctcc agtcacgca ggcgccagg aagcgggtgt gctcgccgat gcgggcgccg 3000
tacttgccct tccactccca gacgcgtcc aggaaggcgg cgcgccgag gtcgtgccgg 3060
ctcttgccct cggctctctt cagctccttc tccaccacca tctgcgtggc gatgccggcg 3120
tggtccgtgc cggggagcca cagggtgttg aagccgctca tccgcttcca gcgcgcgagg 3180
atgtccgga tgggtggcgg gagcgcgtgg ccgatgtgga ggctgcccgt cacgttgggc 3240
ggcggcagca cgatggagaa ggcgggcttg tcggaggtcg cctcggcgcg gaagtagttc 3300
ccctccatcc agcaggcgta ccagcgggcc tcgacctcgg tgggctcgta ggccttggac 3360

agttcagtgg tgtcgggtcat tgcggacggc cagccccggg ggaaccgggg ccggcgggaa 3420
ggctcttgag agaggttaag gcgaggaacg cgtcagtgtc gcgctctccg gtccttgatg 3480
agccgctcga gctcctcacg gatgatggtc tccgccagct gcggtaccac ctcccaggca 3540
atcttctcga tgacctcgcg ggaggccttc gagagcgcct cgcgcagcag cgcctcgcca 3600
ccatccgagg cggtcggacg ggccgtggcc ggagcggcgg tcggcgccac cgcggtcgga 3660
gcggccgtcg gggggccgcc catgtccagg gagatctcct cggcacccgt gggctcgggc 3720
agcgagtctt cgatgtgat ggaaggctgg gctgcctgcg gcgcgcctgt gggagcgccc 3780
aggccgaacg ggtccctcgc gcgggcccgc ggctgggtcg cggcaccggg cgcgggaggc 3840
agcggcgagg cgcgcgccgg aggacgcggg aagcccggct ggctgccctg gggaatcccc 3900
ggacggggca tgccggcgcc ggcccacccg gaggcgcggg cagccccgga cgcggcgcca 3960
tgcccgggcg ggacgcaccc ggccgaggcg gcacgcccgg gacgcgcggc ggcgcccgga 4020
ggcgccatgc cggggcccgg aggacgcggc gcgcccggag caccggcggg cggcattcca 4080
ggccccggcg ggcggggccat gcccggtccc ggcgccattc caggccccgg aggacgcggc 4140
gcgcccggag caccggagg cggcattcca ggccccggag gacggggccat gcccgggccc 4200
ggcgccattc caggccccgg aggacgcggc gcgcccggag caccggcggg catgcccggg 4260
cccggaggac gcggcacgcc cggctgggtg ccctgcggaa tgcccgagc ggcgggccgtg 4320
ggcgcgcccg gctgcggggc cggaggcacc ggcgcgggag tggccggctg ggcggtgggc 4380
gggcgcacct gcgtggccgc ggaggccggc atggtgttgg acttctgccc caccagcgcc 4440
ttcaccttgt cgagcagcac ctggctctcg aagggtctgg cgatgtggtc atcggcgcgg 4500
gcggcgcggg cgcggttctc gtcgaaggcc tcgaagggtc ccgccagcag caccacgggg 4560
atgccctggg tggccgggtc gctcttcagg gcctcgcaga cctcgtagcc gctcttgccc 4620
ggcatcatca catcggcgag gacgacgtcc gggcgcagct cgcgggtacg ggagatggcg 4680
tccagcccgt tgtccaccgc ggtcacctga aagtccctcg tcgcgaagat catcccaatc 4740
accttgcgga tggtagcgga gtcacggcg accagcagat tcttgggcat cggttcgggc 4800
ctcggggggc gtcccacccc cctgaattcg ttggagattc ctgccgcgaa aaaccaggt 4860
cgcggccctg tatcaagcgg ggcagcttac gtttcgcgct ccagcacgat caagaaaaca 4920
tgcgctgcag gtccaggaac agcacggagc ctgtcaaccc gggggcacgg aaggctccaa 4980
tgggtgtctc cggctcgaag tgctgcagca cgcgcagcac ccgggtggcc gtcagcccca 5040
cgttccggcc ggccagctcc gtgaggatga agagggcctc gggtagccgg ggggagccca 5100
gcagcgcggg gaccgagcag atgggcccaca gcacctgggc gtgcgggaag atgcccggca 5160
ccggcccgt ctgcacgggc aggacgtga aggcctcccc tcgcgatata acctgggaga 5220
cgaagcccag cgggacgccc aacaggcgcc cctgcgattc gaagacgagc gccttgctcg 5280

gcgtggactc	ggcccagtg	acggggcgca	gggcccggg	cgcggcacga	gggcccaccc	5340
ggtggggcag	cgcttcggcg	atgagctcca	ggtagagccg	ctccttgtgc	agcaccgcgc	5400
cccggctcag	ggggggccagc	gagtccgcca	gcccggggggg	caggaggaag	aacggcgccc	5460
gcgccacgtc	ggcgacctcc	accacggagc	gcacccgcac	ggccagggtg	ggactcacgt	5520
cgagcaccac	caccatgccc	ggccccctcct	cgggcaagcc	gccgagcagg	gccgacaggt	5580
ccttcacctc	cagcacgcca	cgcaggcttg	agccgttggc	gccgggcatg	gccacttcca	5640
tgacggaggt	ggcctccacc	gcgtagcgcg	tctctcctgc	ctccacgaga	aggcagagcc	5700
ggcgtccgct	ttcgaagg	gtg aca cgg ccc tac cct agc agc ctt cct cgc	5751			
		Val Thr Arg Pro Tyr Pro Ser Ser Leu Pro Arg				
		1 5 10				
ctt ggt ggt aag cca tcg agc cag atg gcg cgc atc ctc ctc gtc gac	5799					
Leu Gly Gly Lys Pro Ser Ser Gln Met Ala Arg Ile Leu Leu Val Asp						
	15 20 25					
gac gaa aag atc gcc cgc acc ctc tac ggc gac tac ctc acg gcc gtg	5847					
Asp Glu Lys Ile Ala Arg Thr Leu Tyr Gly Asp Tyr Leu Thr Ala Val						
	30 35 40					
ggg cac gcc gtc acg gcg gtg ggc tcg ctc cag gac gcg cgg gag gcg	5895					
Gly His Ala Val Thr Ala Val Gly Ser Leu Gln Asp Ala Arg Glu Ala						
	45 50 55					
ctg gcc ggg gac cgc ttc gac gcg gtg gtg acg gac ctc atc ctc ccc	5943					
Leu Ala Gly Asp Arg Phe Asp Ala Val Val Thr Asp Leu Ile Leu Pro						
60 65 70 75						
ggc ggc gac ggc atg gag gtg ctg cgc cac gtg cgc gag cgc cac ccc	5991					
Gly Gly Asp Gly Met Glu Val Leu Arg His Val Arg Glu Arg His Pro						
	80 85 90					
ggc gtg gag gtg gtg gtc atc acc ggc ctg gag aag gtg gac ccc gcc	6039					
Gly Val Glu Val Val Val Ile Thr Gly Leu Glu Lys Val Asp Pro Ala						
	95 100 105					
gtg cgc gcc atc aag agc ggc gcg gcc gag tac ctc gtc aag ccg gtg	6087					
Val Arg Ala Ile Lys Ser Gly Ala Ala Glu Tyr Leu Val Lys Pro Val						
	110 115 120					
gcc ccc gag gcc ctc cag cat gcc gtc cgc cgt gcc ctc acc acg cga	6135					
Ala Pro Glu Ala Leu Gln His Ala Val Arg Arg Ala Leu Thr Thr Arg						
	125 130 135					
gac ctc ctc cag gag aac gcc tcg ctg cgc cgg cac gtc gcc atg ctg	6183					
Asp Leu Leu Gln Glu Asn Ala Ser Leu Arg Arg His Val Ala Met Leu						
140 145 150 155						
gag gcc ggc cag cgc atc gcc acc acg ctg gag cgc gag aag ctg gcc	6231					
Glu Ala Gly Gln Arg Ile Ala Thr Thr Leu Glu Arg Glu Lys Leu Ala						
	160 165 170					
tcc gcc acc acg agc gcg ctg gag tcc atg gcc tgc gcc agc gcc gtg	6279					
Ser Ala Thr Thr Ser Ala Leu Glu Ser Met Ala Cys Ala Ser Ala Val						
	175 180 185					

gtg ctg ctg gag cgc gac ggc gac aag ggc ctg cgg ctc cag ggc gtc	6327
Val Leu Leu Glu Arg Asp Gly Asp Lys Gly Leu Arg Leu Gln Gly Val	
190 195 200	
cgc gga ctg cct ccg cgg cac gag ccg gag ctg gtg gcc acc ctc gtc	6375
Arg Gly Leu Pro Pro Arg His Glu Pro Glu Leu Val Ala Thr Leu Val	
205 210 215	
gag cgc ctc tcg aac gcg cgc gcg ccc cgg gag ctg gaa ggg ctg gac	6423
Glu Arg Leu Ser Asn Ala Arg Ala Pro Arg Glu Leu Glu Gly Leu Asp	
220 225 230 235	
gcc gcc ttc ccg cgc gtc ctc tcc atc ccc gcc gtg gag ggc acc acg	6471
Ala Ala Phe Pro Arg Val Leu Ser Ile Pro Ala Val Glu Gly Thr Thr	
240 245 250	
gtg ctg ggc cac gcg gtg ctc ttc ttc gac ggc gcc gag ccg gag tgg	6519
Val Leu Gly His Ala Val Leu Phe Phe Asp Gly Ala Glu Pro Glu Trp	
255 260 265	
acc ggc gag acg gcg ggc tac ctg gtg cgc aac tgg gca ctg gcg ctg	6567
Thr Gly Glu Thr Ala Gly Tyr Leu Val Arg Asn Trp Ala Leu Ala Leu	
270 275 280	
cgc aac ctc ggg cgc ttc gcg gcg gtg gag gac ctg gcg tac gtc gac	6615
Arg Asn Leu Gly Arg Phe Ala Ala Val Glu Asp Leu Ala Tyr Val Asp	
285 290 295	
gac ctc acc cgc ctc ttc aac acg cgc tac ctg cac ctg gtg ctg gac	6663
Asp Leu Thr Arg Leu Phe Asn Thr Arg Tyr Leu His Leu Val Leu Asp	
300 305 310 315	
cgc gag gtc cag gac tcc gtc cag acg cag cgc ccc ttc agc ctg ctg	6711
Arg Glu Val Gln Asp Ser Val Gln Thr Gln Arg Pro Phe Ser Leu Leu	
320 325 330	
ttc ctg gac ctg gac cac ttc aag tcc atc aac gac acg cac ggc cac	6759
Phe Leu Asp Leu Asp His Phe Lys Ser Ile Asn Asp Thr His Gly His	
335 340 345	
ctc gtg ggc tcg cgg ctg ttg gtg gag acg gcc cgg gtg gtg aag ggc	6807
Leu Val Gly Ser Arg Leu Leu Val Glu Thr Ala Arg Val Val Lys Gly	
350 355 360	
tgc gtg aga gac cac gac gtg gtg gcc cgc tac ggc ggc gac gag tac	6855
Cys Val Arg Asp His Asp Val Val Ala Arg Tyr Gly Gly Asp Glu Tyr	
365 370 375	
gtg gtg ctc ctg cgc aac acc gac tcc ggc ggc gcg ctc aag gtg gcg	6903
Val Val Leu Leu Arg Asn Thr Asp Ser Gly Gly Ala Leu Lys Val Ala	
380 385 390 395	
gag cgc atc cgc cgc acc atg gag acg cac ctg ttc ctc gcg cgc gag	6951
Glu Arg Ile Arg Arg Thr Met Glu Thr His Leu Phe Leu Ala Arg Glu	
400 405 410	
ggg ctg tcg ctg aag ctc agc acg tgc atc ggc gtg gcc agc ttc ccc	6999
Gly Leu Ser Leu Lys Leu Ser Thr Cys Ile Gly Val Ala Ser Phe Pro	
415 420 425	
gag cac gcc cgg gac aag gcc acg ctg ctg gac ctg gcg gac cgg gcc	7047
Glu His Ala Arg Asp Lys Ala Thr Leu Leu Asp Leu Ala Asp Arg Ala	
430 435 440	

atg tac cgg ggc aag cgt ggc tcc cgg aat gtc gtc tac atg gcg gcc	7095
Met Tyr Arg Gly Lys Arg Gly Ser Arg Asn Val Val Tyr Met Ala Ala	
445 450 455	
ctg gat ctg gag ccg ctg ccg gcg gag cgg cgc cag ggc ggc agc ggc	7143
Leu Asp Leu Glu Pro Leu Pro Ala Glu Arg Arg Gln Gly Gly Ser Gly	
460 465 470 475	
ccc ggc ggc tcc agc acg tag cggctcagcg gcgcagggttg cccacggtga	7194
Pro Gly Gly Ser Ser Thr	
480	
ggcgcttggg cgggggtggc tccacctcgg cgcgtgctcc cccgccctcc ttgcccaccg	7254
cctccacctc ctctgtcggg gcaccggtgg tgatgagccg gtactttctcc gaggcctcgc	7314
ggatcatgctg ggctgtctcg gggccaacc gggcaatcag ctgccagaag agggcggcgt	7374
gctcgcgctc ctctgccatg acgtgacgga ggatggcctt ggcctcctcg ttgtcgggtg	7434
catcgatgtg ggccggcgtag aggttgatgg cgtccagctc cgcctcgatg ttcagccgga	7494
tggagcgggc cagctcggag tcggatcatc tccggggcac cagcgagtgg aacgggttgg	7554
tctgcccgcg gaaggccctc ccgaggccgg tgggtggcgtc cacgttcttg acggggcgcc	7614
aggggacgcg aggagcatcc gcgcggccag gggcccggt caacggaatc cctccgccc	7674
cgttgcgtag ataccgctgc ctccgccatg ctccggccat ggaaacaccc gcgcgctct	7734
cccagttgct ccaggccctg gaggcgggag atctgccggc cgcgcgcgcg gcggcggcgg	7794
ctctgcaacg tgcgggcgcg cactcggcac agctggcggc ggaggtgctg cacgagctgc	7854
ggcagcccct gctgggcgtg aaggcctacg cgcagctgct cgcggaggac ggtggcccca	7914
ccggcccgcgt gcggctgctg ctggcgcagg tggagcggat ggaacagatc gtctccgact	7974
acatccgcct cgcagcgcg cggcacgcgc cgcagcagcg cctgtcgtg gcggcgccca	8034
tctgggcggc ggccaagctg ttcagcgtca acccgactc ggcgcgcatg tctctggagg	8094
tggagggcgc cgaggacatc accatccagg gcaacgcgcg gctcatcgag cagctcacgc	8154
tgaacctgct gaacaacgcg cgcgacgcca tggccggccg cgggcgcgtg aaggtggtgc	8214
tctcgcgcga gggcgctct cgggtgctgt acgtggcgga ctggggcccgc ggcattcccgc	8274
tggagctgcg cgagcgcac ttcgagccct acgtcaccgc caacaagcgc ggcacgggac	8334
tggggctggc ggtgtgcaag cgcacgccc aggagcaca cgcgacgatt ggcctcgcgg	8394
cgcggggcgc cattegcgcg gtgccccgc cggccaccgt gttccgggta ctcttccccg	8454
ccacggacgc gccgcgcg ctgcgcaagc ggctgctggt ggtggacgac gagaccatca	8514
tccgcctggt ctccgcgcg ctgatgggca aggagtgcga ggtcatcgag gcggccagcg	8574
gggggggagc gctggatttg ctgcgccagg cgcagtgga cctcatcgct acggacaaga	8634
acgtggggg cctgtccggg ctggagctgg cgcagcaggc gcggcggtc tactccaact	8694

cgcgcgatcat	cctgatgacg	ggctaccogt	cgctgggtgac	gacgcagcag	gcgctggagc	8754
tgggctgtgt	ggactacctg	ctcaagccct	tcgacgacat	ccggcaggtg	cgcggtgtgt	8814
tgcgcacgac	gctgtcctcg	gagccgcggg	tgcctccggg	gggggcccc	ggggccgaag	8874
tgcggcgggt	ggacgtgctg	gaggacaacc	cgacgacggc	gcggtcatc	tccgaggccc	8934
tggacatgct	ggggctggag	gcgcgggtgc	tgcctcgac	ggagctgatg	gcgatggcga	8994
cgccggtggg	cgtgggtggg	agctgggact	tcgcgcgggc	ctatggggcg	aaggcgctgg	9054
agctggccaa	ggcgctggcc	cagggcgccc	ccttcgtcgt	gctggccgag	cacctcacca	9114
tggagacggc	gctggagtcg	ctgcgcggcg	gtgccggcg	gtgccttcg	aagctgctgt	9174
ccgataccac	ggcgctcagc	cgtgagctga	gccgcgcctt	caagagagag	gtcccatgag	9234
actcgcgctc	ctgtccgtcc	tgtcttcg	gggtggcctgt	tcctcggaca	agcccccgcc	9294
gacggagcgg	ccggacgggg	gcgcgacgga	ggactccggg	acgggaggag	aggacgcggg	9354
cccgcgagg	gacggaggca	cggagccgga	tgggtggctcc	ggctccgatg	cgggcacgga	9414
tgcgggagcg	ggcgacgcgg	gcacgggtga	tggcggaagc	tgtgaggcg	ccgcggccga	9474
gcccgatagc	gagctggcca	cgcgactgaa	cacgccccgc	cggtcgcgg	tggacacgac	9534
ggacatctac	atctccgagt	cacactccct	gaacccccag	cagccgagtc	cgggaccggg	9594
acaggttctg	aggctgcccc	gcgcgggagg	tccacccacg	gttctggcca	cgggcttcgg	9654
cgcaccagac	gccatcgccg	tggatgggac	gagcgtctac	gtgctcgacc	tggacgggct	9714
ctggcggtg	gacaaggcaa	ccgggaagcg	gggcgacctg	cccatcgacg	cgacggatcg	9774
caacgtcacc	gtcgggtgta	ccgaggtgct	ccgcgccacg	ctcgccggac	gcgacgtgct	9834
gggtggtcgt	acggggcggc	agtcgctggg	ccgggtggac	accgatggcg	gcaactccca	9894
ggttctctac	accggtcccg	gcggctccct	ggtgcgtggc	gcgcgagtgg	tggactcgga	9954
cgtgtggttc	ctcgtcagcg	cggggacggg	agccggaaac	agctcgggtc	tgtaccgctg	10014
cccgtcggac	ggcagcgcg	ccgcggagcg	gcgggacgcc	accatcctcc	aaggcaattc	10074
actggaggtg	acccccacgc	acttcctcgt	caccgagggc	ggtgggggta	ccggccgggt	10134
gttgcggtg	ccccgcgctg	gaggcaccgc	cgaggtgctg	gccgacggtc	tccagggccc	10194
gtggttcccc	gtggaactga	atgggaccat	ctacttcaag	gagtcacgtg	ccgacggcg	10254
ggacttcctg	cgccgcgtgc	gcacgtgcgc	gctggggacg	tcggatccgg	tggggcccc	10314
gggcacggga	ccaggcgggc	tcacgtgga	cggcagcacc	ctcctctaca	cgtcgcagga	10374
gagcggcacc	ggcggcgccg	tgggcccgtg	gcccgtgacg	cctcctccac	ccccgaaaag	10434
accacggccc	cggcgcctcg	cgaggaggca	tccggggccg	cgtcctgcat	cagtgggaaca	10494
cctcaccggc	tcggcgggcc	cgaggcctca	tcggcccccg	accgccgtga	gcggctcagt	10554
agtccatgtc	gtcgccgcgg	tagtcgggca	tcccggcgcc	gccgccgttc	ttgcccttct	10614

tcttggcggg gcggtcggcg atcatcgcct cgggtggtcag cagcagcgag gccacggacg 10674
cggcgttctg cagcgcggtg cgctcgacct tcgtcgggtc gatgacgccg gccttctcca 10734
ggtcctcgta gacctcggtg cgggcgttgt agccgaacgc gccctggccc tcgcggacct 10794
tgttgatgac caccgcgccc tcgacgccgg cgttggaggc gatcttcgcg agcggtcctt 10854
ggagcgcgcg gcggtatgat tccacgccga agtcctgtc accgccagc ttcagcttct 10914
ccagcgcggg caggcagcgc aggaaggcca cgccgccgcc agggacgatg ccctcctcga 10974
cggccgcgcg agtcgcatgc agcgcgtcct ccacgcgggc cttcttctcc ttcactctcg 11034
tctcggtcgc ggcgcccacg ttgatgacgg ccacgccgcc caccagcttg gccagccgct 11094
cctggagctt ctgcgggtcg tagtcgctgg tgacggtgtc gatctgcgag cggatgagct 11154
tgatcgggcc ctgatctcg ctcttggtgc cggcgccgtc gacgatggtg gtgttgtcct 11214
tgtccaccgt gatgcgcttg gagcgggcca ggtcgttgag cgtcaggttc tcgtacttgt 11274
ggcccagctc ctgcgtgacg accatgccgc ccgtcagggt ggcgatgtcc ttcagcatct 11334
ccttgcggcg gtcaccgaag cccggcgctt tcaccgcggc cacgttcagc acgccgcgga 11394
tcttgttgac caccaggggt gccacgcct cgccctcgat gtcgtcggcg atgatgagca 11454
gcggcttccc ggagcgcgcc acctgctcca ggatgggaat catgtcctgc atcgacgaga 11514
ccttcttctc gctgatgagg atgaaggggt cgtccatgac gacctccatg cgctcgcggt 11574
tcgtcacgaa gtacggagag acgtagccgc ggtcgaaactg catgccctcc accacgtcga 11634
ggttggtctc caggcccttg gcctcctcga cggatgatgac gccctccttg cccaccttct 11694
ccatcgcgtc cgcgatgatg gagccgatgg tctcatcacc gttggcgag atggtgcccc 11754
cctgggtgat ggccttcttg tccgcggtgg gcttggacag cttcttcagc tcctccacga 11814
cgacctcgac ggccttgctg atgcgcgct tgagggtccat ggggctgtgg ccggcggcca 11874
ccagcttgag gccctcctcg tagatggcac gcgccagcac cgtggcggtc gtggtgccgt 11934
cgccggcctt gtcggaggtc ttcgacgcga cctccttcac catctgcgcg cccatgttct 11994
cgaacttggt gtcgaggctg atctccttgg cgacgggtgac gccgtccttg gtgatcgtgg 12054
gcgaacccaa gctcttctcg atgaccacgt tgcggccctt ggggccagg gtcaccgcca 12114
cggcatccgc gaggttccgg acaccgcgaa ggatggcctc acgcgcggac tgggtggaaga 12174
aaatctcctt cgctgccatc gcaacctcct gaaattactg ggggatttaa ctgtgacgac 12234
gctggccgtt agcactcggc catggcgagc gccaacataa gggccagcga aatgatgtca 12294
accaggaagg gcgtgcctgt gggcgagcgg acggggccag ggccctcgg ggccgtctgg 12354
ccggggggct actcggcgaa gcgcacgagg ttgagcagct gacctatgtc caggggtttg 12414
cgccagctca gcgccacgcc cttgcgcagg ccccggtcgg tgacgcctc gtctgggcg 12474

cccgtcatca gcagcacggg aatggcctgg aagcgggggt cggccttgag catctccgtg 12534
 aagtggatgc cgtccaggtg gggcatgagg taatcggtga tgaccatgcc cacctggttg 12594
 cgctccagga tgtcgagcgc ctgcagggca tccgccgccg agtagaccgt gtagccatgc 12654
 atctccagga agcgagagag catggtgcac agctcgagat cgtcatcgac gacaaggatg 12714
 ttcacggggc gggccgtcac acatggacgg gcgccggaac ctaacagccg tgcccgtcgt 12774
 tgacaacgca cagcgtgcct tcatatgtgc cggcccatgg ggaagcggac cgcgaagcga 12834
 gagggg gtg ggt ggc gtg gag gct cgg ctg gat gcg gtg gcg gac gct 12882
 Val Gly Gly Val Glu Ala Arg Leu Asp Ala Val Ala Asp Ala
 485 490 495
 ttc gag gcc ggg gac ttc gag gcc gcg ctg gcc ggg gcg gag ggc ctg 12930
 Phe Glu Ala Gly Asp Phe Glu Ala Ala Leu Ala Gly Ala Glu Gly Leu
 500 505 510
 ctg gcg gac gcg ccc gag ctt ccc gag gcc ctc cac ttc cgg gct tcg 12978
 Leu Ala Asp Ala Pro Glu Leu Pro Glu Ala Leu His Phe Arg Ala Ser
 515 520 525
 gcc ctg gtg gaa ttg ggg cgg ctg gag gag gcc ggg aag gcc ttc ggg 13026
 Ala Leu Val Glu Leu Gly Arg Leu Glu Glu Ala Gly Lys Ala Phe Gly
 530 535 540
 cag gcc ctg aag gtg gcg ccg gag gac ctg gaa atc ctg ctg agc gcc 13074
 Gln Ala Leu Lys Val Ala Pro Glu Asp Leu Glu Ile Leu Leu Ser Ala
 545 550 555
 gcg gac tgc ctg gtg tgc cgc gcc ggg gag gac cgg gag gcg gtg gcg 13122
 Ala Asp Cys Leu Val Cys Arg Ala Gly Glu Asp Arg Glu Ala Val Ala
 560 565 570 575
 gag ggg ctc gcc ctg tgc gcg cga ggc cgg cgc ctg gcc cag aag gcc 13170
 Glu Gly Leu Ala Leu Cys Ala Arg Gly Arg Arg Leu Ala Gln Lys Ala
 580 585 590
 gac gac gtg gag atg ctg tac gag ttc ctc ctc ctg gag ggc atg ggg 13218
 Asp Asp Val Glu Met Leu Tyr Glu Phe Leu Leu Leu Glu Gly Met Gly
 595 600 605
 ctc aac cag atg ggc gag tgc gcc acg gcg ctg gtg agc ctg gac gcg 13266
 Leu Asn Gln Met Gly Glu Cys Ala Thr Ala Leu Val Ser Leu Asp Ala
 610 615 620
 gcg ctc ggc cac atg ccg cgc tcg ctg gac gcg cag gtc gag cgg ggc 13314
 Ala Leu Gly His Met Pro Arg Ser Leu Asp Ala Gln Val Glu Arg Gly
 625 630 635
 atc gcg ctg ttc gag ctg tgc cgc ttc gac gag gcg aag gcc gct ttc 13362
 Ile Ala Leu Phe Glu Leu Cys Arg Phe Asp Glu Ala Lys Ala Ala Phe
 640 645 650 655
 gag aag gtg ctg aag gac gcg ccg gac gac ccg tgg gcg cac cac tac 13410
 Glu Lys Val Leu Lys Asp Ala Pro Asp Asp Pro Trp Ala His His Tyr
 660 665 670
 ctg ggg ctc atc gcc gag cgg cgc ggg gac gag aag gag gcg aag cgg 13458
 Leu Gly Leu Ile Ala Glu Arg Arg Gly Asp Glu Lys Glu Ala Lys Arg
 675 680 685

cgc ttc gac aag gcg cgg gcg ctg gtg ccc gag gag ttc ccc ccg ccg	13506
Arg Phe Asp Lys Ala Arg Ala Leu Val Pro Glu Glu Phe Pro Pro Pro	
690 695 700	
gtg gag ctg gcc gag gcg gag ttc gac cgc gcg gtg gag gac gcg gtg	13554
Val Glu Leu Ala Glu Ala Glu Phe Asp Arg Ala Val Glu Asp Ala Val	
705 710 715	
aag tcc ctg ccc cgc cac gcg aag cag tac ctg gac aac gtc acc atc	13602
Lys Ser Leu Pro Arg His Ala Lys Gln Tyr Leu Asp Asn Val Thr Ile	
720 725 730 735	
gcc gtg gag gac ctg ccc tcg gac gag gat ttg ctg ggg cag gac ccg	13650
Ala Val Glu Asp Leu Pro Ser Asp Glu Asp Leu Leu Gly Gln Asp Pro	
740 745 750	
ccg ctg tcc ccg agc atc ctc ggg gtg ttc cgc ggc acg ccc gtg ggc	13698
Pro Leu Ser Pro Ser Ile Leu Gly Val Phe Arg Gly Thr Pro Val Gly	
755 760 765	
gag cgg agc gtg atg aac gcg tat gag ctc ccc gcg tcc atc gtg ctc	13746
Glu Arg Ser Val Met Asn Ala Tyr Glu Leu Pro Ala Ser Ile Val Leu	
770 775 780	
tac cag cgc aac ctg gag cgc ttc gcg agg acg cgc gag gaa ctc atc	13794
Tyr Gln Arg Asn Leu Glu Arg Phe Ala Arg Thr Arg Glu Glu Leu Ile	
785 790 795	
gag cag att gga atc acc gtg atg cac gag gtc ggt cac ctc atg ggg	13842
Glu Gln Ile Gly Ile Thr Val Met His Glu Val Gly His Leu Met Gly	
800 805 810 815	
ctg gat gag gac gac ctg tgg cag cgg ggg ctg gac tga cgagcgaggc	13891
Leu Asp Glu Asp Asp Leu Trp Gln Arg Gly Leu Asp	
820 825	
ccgcgtggct tccagcgtct ggcggatgcc ctcgctgtag ggcgtcttgc gcacgctgcc	13951
cagcagcttg tgcagcgccg agtcgtccat gaggacgggc tcggtgagca ggtagtgcac	14011
ctcgaccatc tcgcgcatga acgggttgaa gaggccgatg aggcggagca tccccttccc	14071
cgccgtcatg aacttcgga ggtggccggc ctgggcgtag atctctcga cgagcttgcg	14131
ctgtgaggtg acgccggcgc cggcgaggtt ccaccagcgg ccgtaggcac cgggggcgtc	14191
catgagggcc gtcacggtgg gaccacgctc gggcacgtag acgtactcgt gcggcgggtc	14251
gatggggccg atgagctggg cgcgcttgcc ctgggcggcg gcgacgaagg cggagtggag	14311
gaagctcttg tcgatgccgg ggccgtagaa gtccgggagg cggaggatgg tggcgcgag	14371
ctttccggcg gcgtccgcg cgaggaggat gtcctcctgc tccttgcgca tggggccctt	14431
gaaggtgtgc ggctcgcggg ggtggtcctc ggtgacgggc gtggtgcggg ggcggccgta	14491
cgggtacacg gtgccgatga gcaccacgcg ctcgacgccc tcggcgatgg cggcgtccac	14551
ggtgcggcgc atgagctccg ggtggagctg gaagcgccag tagtccacgc cgaccatgta	14611
gatgagcgtg cggatgcccc gggcggcggc gcggatggtg tccggcgcggt cggggttcca	14671

gggtggcgatc	tccgccttcg	gggtccgcgcc	gaactgcgac	tgaagcgagg	cgcgcgagcg	14731
gcccaaccacc	cggtacgccc	tcccctgctc	gcgaagagcc	ttcacgatgc	tgtgtccgat	14791
ggggccccgac	gctccaaaca	acgccacctt	gtccgtatcc	atgggtccgct	ccctggggttc	14851
cgaacggcgga	gcgttttatg	aacgtcgttc	attgtgaacg	gtgttcacat	actgctggcc	14911
gttcaggacg	tcaagcccta	cagtggggat	atg ggg att tcg gag cgg aag gag			14965
			Met Gly Ile Ser Glu Arg Lys Glu			
			830		835	
cgc cag cgg gcg gag ctg cgc gag cag atc ctc cag gtg gcg aag gac						15013
Arg Gln Arg Ala Glu Leu Arg Glu Gln Ile Leu Gln Val Ala Lys Asp						
	840		845		850	
atg gtg acg cga gag ggc ttt ggt gcc ctc tcg atg cgc aag ctg gcg						15061
Met Val Thr Arg Glu Gly Phe Gly Ala Leu Ser Met Arg Lys Leu Ala						
	855		860		865	
gac gcg gtg gag tac gcg ccg gcg acg ctc tac ctg cac ttc gag aac						15109
Asp Ala Val Glu Tyr Ala Pro Ala Thr Leu Tyr Leu His Phe Glu Asn						
	870		875		880	
cgg gac gcg att gcc ccg gag ctg tgc atc cgg ggg ttc cag gac ctg						15157
Arg Asp Ala Ile Ala Arg Glu Leu Cys Ile Arg Gly Phe Gln Asp Leu						
	885		890		895	
ctg gag gcg ttc gag ccg gcg gca tcc gtg gag gag ccg gtg gag cgg						15205
Leu Glu Ala Phe Glu Pro Ala Ala Ser Val Glu Glu Pro Val Glu Arg						
900		905		910		915
ctg tac cgg ctg ggc gag gcc tac gtg aag ttc ggg ctg gag cag ccg						15253
Leu Tyr Arg Leu Gly Glu Ala Tyr Val Lys Phe Gly Leu Glu Gln Pro						
	920		925		930	
gag acg tac cgg ctc atc ttc atg gag gac ccg aag ctg tcg acg gcg						15301
Glu Thr Tyr Arg Leu Ile Phe Met Glu Asp Pro Lys Leu Ser Thr Ala						
	935		940		945	
ctg ttc agg gac gcg ccg gag gac ggc gcg ggg ccg cgg tcc ttc ggg						15349
Leu Phe Arg Asp Ala Pro Glu Asp Gly Ala Gly Pro Arg Ser Phe Gly						
	950		955		960	
gtg ctg gtg aag gtg ttc gag gac ctg aag gcg gcg ggg agg att gcc						15397
Val Leu Val Lys Val Phe Glu Asp Leu Lys Ala Ala Gly Arg Ile Ala						
	965		970		975	
gag gac gcc gag ccc tcg aag ctg gcc gag gtg ctg tgg gcc ggg gtg						15445
Glu Asp Ala Glu Pro Ser Lys Leu Ala Glu Val Leu Trp Ala Gly Val						
980		985		990		995
cac ggc atc gtg gcg ctg ccg ctg acg tgt acg ggc ttc aag ggc						15490
His Gly Ile Val Ala Leu Arg Leu Thr Cys Thr Gly Phe Lys Gly						
	1000		1005		1010	
tcg ccg ccg gag gag ctg gcg ccg ctg ttg gtg tcc acg atg gtg						15535
Ser Pro Pro Glu Glu Leu Ala Arg Leu Leu Val Ser Thr Met Val						
	1015		1020		1025	
aac gga ttg ccg ggg ttg aag ccc gtg aag gcg gcg ccg aag aca						15580
Asn Gly Leu Pro Gly Leu Lys Pro Val Lys Ala Ala Arg Lys Thr						
	1030		1035		1040	

cgg tga gccgggcggtt gggtgacggg gcccggtggg tgccggcagta atcgacgggtg 15636
Arg

gtagggcctc cggacctggt ggccggccccg gtcttcaaaa ccggtgaggg gcgctgagag 15696

gcgtcccttg tgagttcgat tctcatgctc taccgccatg ttttccttgg ggtttttctga 15756

cgcgggccctg atcggttgc aggacgcccc acccgctgt acctctcgta cccaagtttt 15816

acccaagtga ggggggtacg ggagccgcgc atg acg aag tgg agc ggg aag tgg 15870
Met Thr Lys Trp Ser Gly Lys Trp
1045

atc gga ggc cgc acg tac acg gcc acc gac gga agc acg cgc tgg 15915
Ile Gly Gly Arg Thr Tyr Thr Ala Thr Asp Gly Ser Thr Arg Trp
1050 1055 1060

atc atc cgc aag acc gtc gcc ggg gtc gcg tac aac gtc acg ctg 15960
Ile Ile Arg Lys Thr Val Ala Gly Val Ala Tyr Asn Val Thr Leu
1065 1070 1075

gac gtg cgg agc gag gcc gag gcc ctg gct gag ctg gcg gcc ttc 16005
Asp Val Arg Ser Glu Ala Glu Ala Leu Ala Glu Leu Ala Ala Phe
1080 1085 1090

cgc cgc aac ccg gcc gcg tac cgg acc gcg agc cag gag cgc caa 16050
Arg Arg Asn Pro Ala Ala Tyr Arg Thr Ala Ser Gln Glu Arg Gln
1095 1100 1105

gac gag gcc caa cag gcc cag gag gac gcc gcg cag gcg gtg ttc 16095
Asp Glu Ala Gln Gln Ala Gln Glu Asp Ala Ala Gln Ala Val Phe
1110 1115 1120

ctg gac gag gat cgt gcc cgg cgg ttc ctc cag cat ctg aag tcc 16140
Leu Asp Glu Asp Arg Ala Arg Arg Phe Leu Gln His Leu Lys Ser
1125 1130 1135

agc ggc agg tcc gac gag tac ctg aag agc acg cgc agc tac ctg 16185
Ser Gly Arg Ser Asp Glu Tyr Leu Lys Ser Thr Arg Ser Tyr Leu
1140 1145 1150

gcc gcg tgg gca acg gcg ctc gcg gac aag gac ctg cgc gcg gtg 16230
Ala Ala Trp Ala Thr Ala Leu Ala Asp Lys Asp Leu Arg Ala Val
1155 1160 1165

acc gga ccg gcg ttg aag aag atc ctg gcc acc tgg gac acc ggc 16275
Thr Gly Pro Ala Leu Lys Lys Ile Leu Ala Thr Trp Asp Thr Gly
1170 1175 1180

aag aag tgg agg atc atc gtc ttc aag tcc ttc acc tcg ttc ctt 16320
Lys Lys Trp Arg Ile Ile Val Phe Lys Ser Phe Thr Ser Phe Leu
1185 1190 1195

cag gac tcg ggc gaa ctg gac ccg atg gtg aac ccc gga cgc ttc 16365
Gln Asp Ser Gly Glu Leu Asp Pro Met Val Asn Pro Gly Arg Phe
1200 1205 1210

aat aag gtc ccc cct gcc cgc cgg aac cac gaa ctc aag ggc tac 16410
Leu Lys Val Pro Pro Ala Arg Arg Asn His Glu Leu Lys Gly Tyr
1215 1220 1225

gccttccccg agactcgcgc cgcaattgcc ggagccaagc cgaccagagc ccccatcacg 17413
gttgcgaggat tcggaggcgt gctgcgctcg ctgctccctt ccacgcccac tcaccccgtc 17473
gggactattc ccgcgtccat tccgatcgac gccatcaaca tcgttgcggc ggacctgtac 17533
ccgaaccgga cctgcgatag ggagcacttg gaagagttga cgaagggtgct caacacggag 17593
ttggaggcgc agtaccgcac gcgcctggag tggtcaggag acacgctgca cctgcgcgcc 17653
aagtagacca tcggcccctc gggcgcggcc agcaaagtct ctcccttgat cgaggagtag 17713
gatgcggcga actacgcctc gatcactgcc ggatcattct tgtcacgatg tcgggggggtt 17773
agagcgtgtc gatccggctg cggatgccag gacaagaagt tctgcctcct ctatcccggt 17833
gggccatgtc gaggggcacc aaaaccctcg ctcgtagagct tcggccgctt cgttacggcg 17893
actttcaacg aggctcaaat cgtgagcttg cggatggcaa cacctgggtt cggatacgct 17953
ccggtgcgg ttccctcctg ctcccccgc cgctccttgg tgcggactcg ttcaggctgc 18013
ttggtttgtt gaggccgatt gcctcccttc cgaggggccc cctccgcgag cagtgcctcg 18073
atgctcgcca ccgtgatcat cacggaccgt ccgatcttcg gtgcccgtc caaccggcct 18133
tgctccagga gcttgaagac ctgtgaccgc ttgcagccga gcattgccga cgcgcgctcc 18193
agcatcacag ccgaggcggt ctgtgtgacc gattgagcgc gaagcgcacg aacctcgcgc 18253
ttgatctcgg ccagttcctg gaggatggcg cattccacgg gtgcgggctc gtggctcctca 18313
atgggcacgg caggtcgaaa gactggaatg ggaataggaa tgagggtcaa cgggtggctcc 18373
ggccgagcaa gcaaggagcc tgtacggata tagctgaagt tctcgatcat ggacttgtac 18433
tcaaccggtc tccgtcggag gccgtggatg gtgtcgaagg tgtggatcga gccggtggcc 18493
accagtccgt ccctattctg aatgtatcga atctctccga ggctccttcg ggcgctcgcc 18553
ctgcccgcct ccgtgccggt gaaggcgaac gggaggcgct tggcccggca cctgagcccc 18613
gttggaacgc atggatggcc cggaagggtc acagcacggc ctacttcagc cctttactcg 18673
tgctgcggcg tcatcgctct gacttcccca tagccttcca tcccttcga tccctccacc 18733
tctctctgca accaggggga gaccgaagca gacagcatgg aaggcatcgc cgcagtcgga 18793
acgcatccgt tccaatcacg ggtgcgcaag gtccatccga gcgccagcgg cgatgctctc 18853
catcaccccc tcgatctcgg tctcccagcg ttccaccgtg acccggtact tgattccatc 18913
gctccgcttc cagcccggtg tctcttcacc ggccaactcc ttcattgcgtt caccgaactt 18973
cgtgagcgag aacgccatcc ggtcacggcc ctggtacttc aagtcctcga cgtagtggcc 19033
gtacagcgcc gtggccttga tccactggtc ccgtgcgcg caccggcgct cgaagaaatc 19093
ttccaccgag tccctggctc gaatctcatc ctggcgggcc tggacctgcg ccaattggcc 19153
gaggagcggg ctctccgct tctcgtccac gctcgtccac atggccagga agtcgatctg 19213

ctccagtgcc gtccagttca tctggggcag acagtccagt tgatagaagc gtcgcattcc 19273
cgtgggatcg acgacgatct ccttcacctg cttgttggaac gcaccgatga aggtggcggt 19333
gttggtgccg gtctggagct tctggacacc gagcaccggt tactcgatcc gctcggcggt 19393
gatgcggttc ttgagcacgt cagacgctac ccgctctgcc ttcgccatth cgtcgaagaa 19453
catgccgtag cagcgagtca accggaactg ctggcgctca tcggtcagga actggagatc 19513
ggcggggaac ccaacgagat cgccaagagg ctgtagcagg ctcttgatgg ccgtagtctt 19573
cccgcgccgc tgtttgccga cgaagatcgg catcaggtgg tccttgaccg gcttcccga 19633
caacttcgc ttcacctgcc aatgaagtg aagaacgacc acgagatcta ccgggtcaag 19693
gcgtcccgtg gttgctcgaa tgaactctc ggcgctacct cgcgctccgt cctcgcgga 19753
cttcaacttc tcacgatact gctcgacgat ccccgggcgt tggctccgct ccacaggcg 19813
cacggcgta ccgagatttc gcacgctgat gcatgtgcgc tcgttcgtgg actcgatctc 19873
gatctgggac aagaactcgg acgcaggaaa catccggccc ccaatctgga agccatcgct 19933
cgcgtaacg aatgtgacgc cgttcttctc catccagtgt tgaacccatg tgcgcggctc 19993
gaagcgggttc tcgaccctcg gcaacgcctg caccaatcgg tcgatgacct cgggcttgca 20053
gaccttcttc ctcttcaagg tcccgacca ctcttgaag atgccatgct cggtggagaa 20113
ccaaaccgag caccagacga gagcgcgctt gatctcgatc tccttcaccc caccgtgcgc 20173
ctcatgctgt tccacgagcc agtcacagag tgccgaagga agcgccgggc gcgcagaggc 20233
caagtactgc acgatcagga actgacgagc ccatgcttcc gtcttggtgt gatgagcggc 20293
aagcaacttc ttcattctgc gatgcacca atcatccaca ccggatttct cgtgctgtgg 20353
gtcgtggtcc acgtacggaa tctgagggtc acatccgcgc gcagaaaggc ccttcgcgat 20413
ccgcacaact tcacgaacga cgttcatggt cgtgcgctcc atgtccggcg catcgccgca 20473
gacgaacacc gtgcggccct gtgcaaggta tgatgcggcc tcggggtgcg gtaccactc 20533
gtccgtcagg tcatcccagc ttcgtcgccg ggcctcaaca tcatgggcca tgctgatgcc 20593
ggggagcccc agcgccgggg tgcttctctg cgcgagcagc gccgccttct tctcgccctc 20653
aatgagcgcc atggggactt ctctgtcttg aagcatcggg cagaggatcg cggtgtcgaa 20713
gatctggggc gcgtacccga aaggagcctc gtacttcttc accttcgctg gctccccttc 20773
gacctatcc tgggtgtacg ggcgaggat gtggtccgtg aaggctcgcg ttgggagggt 20833
gagctgcgcc tccttcggcc ggagcacgcg cggcgatgc ggcttgacct tggcgtagtt 20893
gagccccggg tacacgagca ccagccctcc agcgtcccc ttccattcct cgggcccagtt 20953
caacaacttc cggttggtct cgtccgtgtc cgttcgcaa ccgagtttcc gcaccgtcga 21013
gggcaggagg caggacttct tcaggctctc gcagtctgc gccgagatcg gaagctcgtt 21073
ccactcagcc tcgttgcggc tggattcaac ctcgttcgcc ttcgagtcgg gcagtgcctt 21133

gtccgagtc	tcattgtccca	ttgggtacac	ctctcccatc	aacgccgaag	ccgtgcttcg	21193
acgccttgaa	atgcaggggtg	gctcttgaac	ttgtatogat	tctggctacg	tcgttgccggc	21253
gccctggagc	ggagccgggt	tcgggtgtggc	tacaaggaat	atatcgaatg	tgtctgcgca	21313
cctcggacgg	tgcgccccga	gggccacggg	catcaacgcc	gctgcgcgat	cctgcgccgt	21373
ggacggccag	gggattcgcg	acgtacaccg	gacggcaacc	ttccgtgtat	gagccgaggc	21433
acccgctgga	gcacggagta	cctggacgcg	atgctcgcgg	cctgggagag	gtccgggctg	21493
acctcggagc	agttcacggc	cacggccgcg	ccagggctgc	ctgcgcgtac	cctacggctg	21553
caccgccagc	gaaaagagaa	ggagatccac	gcagagcttc	agcggcttcg	ggatgaaaac	21613
gcgacgctcc	ggtcgatgct	cgaaaggagc	gcgttgggtcc	ctgggtggttg	ccggggcggca	21673
accgtcagcg	acccaacgg	ccacggcatt	ggagatcgcc	atgtggcaaa	cggaaccgaa	21733
gccccttcgc	cgaccgcgac	cggcccgcga	gcattggagc	cggtacagcc	gcgcgtggtg	21793
ccgcaggacc	aggttgccgg	accctcccc	ggtgatcacg	atggctcgtgt	gctcgtaggg	21853
gaaatcaccg	agccggtgcc	cccgaagaga	ccggcgttct	cctgggattg	atggcctcgc	21913
gcctgctggc	cgatgtcggg	cgccatctac	cgctgggtctt	gtcccgcgag	cgcccttcaag	21973
agcttcttga	tctcggccga	ggacatcgtg	cggatgcggt	ccacgaggcg	ttgccgctcg	22033
tcctcctcgg	gaaacgtcag	gagatccagc	ggcaacactc	ccaggtggtc	cgccagcacc	22093
ttcaacgtgc	gaaggccaag	ctcctgcggc	cagtgccttag	aggtgggact	cagggcacgg	22153
agcagaggca	gagcgcgagc	accaggcagg	cgctccgccg	cgctcaccct	gccgagccct	22213
gtggtcgcctt	gtaataccct	ttgggtttat	gaatagtaga	cctcggaccg	cgagagccac	22273
cggccccgtc	cccctcagcc	gaggccccac	cagggagccg	gaaggcccga	gccagtcccc	22333
gccgcgcgtc	ccttcccacc	tggaggccgc	ac atg acc gat gac agc acg ccg			22386
			Met Thr Asp Asp Ser Thr Pro			
			1420		1425	
acc gac gaa acg cag atc gag gcc ccg acg ccg gag aac acc tcc						22431
Thr Asp Glu Thr Gln Ile Glu Ala Pro Thr Pro Glu Asn Thr Ser						
	1430		1435		1440	
gac ctg aag ggc gcc cgc ccc cga ctc cgg ttc acc atc gag acc						22476
Asp Leu Lys Gly Ala Arg Pro Arg Leu Arg Phe Thr Ile Glu Thr						
	1445		1450		1455	
acc acg tac gag acc ctc cgc acg ctg gag cgg cgg ggt aac ggc						22521
Thr Thr Tyr Glu Thr Leu Arg Thr Leu Glu Arg Arg Gly Asn Gly						
	1460		1465		1470	
gag gtg gtg ctc ctg gcc gag cgt cac ctc ccc cat gga ctc gcc						22566
Glu Val Val Leu Leu Ala Glu Arg His Leu Pro His Gly Leu Ala						
	1475		1480		1485	
ggg ctc gtc acc atc aag cgc ctg cgc aat ccc gtc act ttc gag						22611
Gly Leu Val Thr Ile Lys Arg Leu Arg Asn Pro Val Thr Phe Glu						

1490	1495	1500	
cgc tgc cag cgg Arg Cys Gln Arg 1505	ctg atc gag gag gtc Leu Ile Glu Glu Val 1510	cag ctc tcc ttc cgc ctc Gln Leu Ser Phe Arg Leu 1515	22656
cac cac ccg gcc His His Pro Ala 1520	atc gcc cag gtc cat Ile Ala Gln Val His 1525	cac ctg aaa atc cat gcc His Leu Lys Ile His Ala 1530	22701
gac cgg ccg cac Asp Arg Pro His 1535	gtc atc gcc gag tac Val Ile Ala Glu Tyr 1540	gtg gac ggc ccc acg ctg Val Asp Gly Pro Thr Leu 1545	22746
gac acc atc atc Asp Thr Ile Ile 1550	agc ctc gcc acc atg Ser Leu Ala Thr Met 1555	cgc gag aag ccg ctc tcc Arg Glu Lys Pro Leu Ser 1560	22791
gcg ccc ttc gcc Ala Pro Phe Ala 1565	ctc tac atc gcc gcc Leu Tyr Ile Ala Ala 1570	gag gtg gcc gac gcc ctc Glu Val Ala Asp Ala Leu 1575	22836
cac cac gcg cac His His Ala His 1580	acg ctg aga gat tcg Thr Leu Arg Asp Ser 1585	gag aac cgg ccg ctg ggc Glu Asn Arg Pro Leu Gly 1590	22881
atc atc cac cgc Ile Ile His Arg 1595	gac gtg gcg ccc cgg Asp Val Ala Pro Arg 1600	aac atc cgc gtg gcc cgc Asn Ile Arg Val Ala Arg 1605	22926
agc ggc gag gtg Ser Gly Glu Val 1610	aag gtg acg gac ttc Lys Val Thr Asp Phe 1615	ggc gct acc tac tcg ctc Gly Ala Thr Tyr Ser Leu 1620	22971
atg gtg ggc agg Met Val Gly Arg 1625	gag gag acg ccg ggc Glu Glu Thr Pro Gly 1630	ctc ctg ctc aag ggc gac Leu Leu Leu Lys Gly Asp 1635	23016
gtg gcg tac gcc Val Ala Tyr Ala 1640	tcg ccc gag tac ctg Ser Pro Glu Tyr Leu 1645	aac cgt aag ccc atg gat Asn Arg Lys Pro Met Asp 1650	23061
ggc cgg tcc gac Gly Arg Ser Asp 1655	atc ttc tcg ctg ggc Ile Phe Ser Leu Gly 1660	ctc gtc ctc atg gag atg Leu Val Leu Met Glu Met 1665	23106
ctg acg tgc aag Leu Thr Cys Lys 1670	cac ctc ttc gac gtg His Leu Phe Asp Val 1675	gag gac gaa aag gcc ccc Glu Asp Glu Lys Ala Pro 1680	23151
aac gcc acc gtg Asn Ala Thr Val 1685	gac gtg aag acg gag Asp Val Lys Thr Glu 1690	gag acg ccc tcg gtg ccc Glu Thr Pro Ser Val Pro 1695	23196
ctc acg cag atg Leu Thr Gln Met 1700	atc gcg ctc gtc aac Ile Ala Leu Val Asn 1705	cgc tac cgc ccc gag gac Arg Tyr Arg Pro Glu Asp 1710	23241
gtg gag aac gcg Val Glu Asn Ala 1715	atg gcg ggc ctg ccg Met Ala Gly Leu Pro 1720	gac gcg ctc aag gcc atc Asp Ala Leu Lys Ala Ile 1725	23286
gtc cac aag gcc 1730	ctc cag cga aag ctc 1735	tcc gag cgc tac gcc acg 1740	23331

Val	His	Lys	Ala	Leu	Gln	Arg	Lys	Leu	Ser	Glu	Arg	Tyr	Ala	Thr	
			1730					1735						1740	
gcc	gcc	gag	atg	cgc	gat	gcg	ctg	cga	gcg	gcg	ctc	gcg	gcg	gag	23376
Ala	Ala	Glu	Met	Arg	Asp	Ala	Leu	Arg	Ala	Ala	Leu	Ala	Ala	Glu	
			1745					1750						1755	
tca	cag	ccc	ttc	ggc	cgg	aag	gag	gcg	agc	gag	gag	ctg	gcg	cgg	23421
Ser	Gln	Pro	Phe	Gly	Arg	Lys	Glu	Ala	Ser	Glu	Glu	Leu	Ala	Arg	
			1760					1765						1770	
atg	ctg	tcg	gag	gcg	tcc	gtc	ctg	cgc	gac	cgg	gtg	gaa	ctg	gac	23466
Met	Leu	Ser	Glu	Ala	Ser	Val	Leu	Arg	Asp	Arg	Val	Glu	Leu	Asp	
			1775					1780						1785	
gaa	gaa	ggc	atc	ttc	ccc	gag	ggc	ctg	gac	gcc	gac	gag	ccg	acg	23511
Glu	Glu	Gly	Ile	Phe	Pro	Glu	Gly	Leu	Asp	Ala	Asp	Glu	Pro	Thr	
			1790					1795						1800	
ccc	gca	ccg	aac	gaa	gag	tga	gaacggg	cgg	cggtgcattc	cttgaagcaa					23562
Pro	Ala	Pro	Asn	Glu	Glu										
			1805												
gcggggccccg	gtgtgggccc	gcgtcgtttc	agcgcaacaa	cacctggggcc	accttcacga										23622
gggcatcgag	ctggtcctcg	tccatcttgc	gcgccagtc	ggtgagctgg	cgcagcttgg										23682
gcgccccgcc	ctgccctcgc	tcccccttgc	ccgacttcgt	ccccctctcc	gcgtccgcga										23742
tccccagcaa	ctcgtccgag	gagatgcgca	gcaccgagca	catccgcagc	agcgtctgga										23802
cgctgggcag	catcttccccg	cgtccaggc	ggctgtagac	catgtgcgcg	aggccagct										23862
tctcgccac	ctccgcctgc	gtgagcccga	gctgcgtccg	ggcctcacgg	gcggcacttc										23922
caatacgggt	cgccagctct	tcattcatgc	gtcggggtac	caggacaccg	agaaccagg										23982
ttggcgcggg	caggaaggcc	caagcctcgg	ggacgcccct	tcctgcccga	tctcggcggc										24042
catgtcgatc	cacgagcagg	gccccgtcac	gcggcgcccc	tattggccgg	acccatcggg										24102
ggcgggaacc	tccagggcca	ggccgcccag	gatgaagtcc	acggggacac	ggaggacacc										24162
gcacaggccc	accagcgtgg	ggaggctggg	cagcaaccgg	ccacgctcca	ggcggctgaa										24222
caccagcgtc	ggcacgtgga	tggcctcggc	cacctggggc	tgcgacaggc	ccgcccgatg										24282
cggggcgggc	cggacattgg	ttccgatctg	ctgggagagc	ttcgcgtaca	tgaggagacc										24342
tgggagaaa	aggagaatgc	cgcgtggggc	gcaactccgag	catacagctc	accaggttca										24402
accatagtat	ccgctcaagg	tttacccttc	tcccgcgcca	cttcgcgggg	ggaactcgta										24462
cgagatccgc	taccgggtga	gagcatgtcc	aggaacgccg	acccacacg	ccccgcggaa										24522
ggagaggaca	cttgagcccc	cgtctgcattc	ccgtcgtccc	ccggggcacg	gacatcgggc										24582
gagacgctg	gtg	gag	gag	aag	ctg	gga	gcc	ggg	ggc	ttc	ggt	gcc	gtg		24629
	Val	Glu	Glu	Lys	Leu	Gly	Ala	Gly	Gly	Phe	Gly	Ala	Val		
			1810					1815					1820		
gag	ggc	ggc	cgg	cgt	gga	ggg	cgc	tac	gcg	ctc	aag	ctc	atc		24674
Tyr	Arg	Ala	Arg	Arg	Gly	Gly	Arg	Leu	Tyr	Ala	Leu	Lys	Leu	Ile	

cca ctc tgg ggg ctg	gct gag tgg gcc gag	cgc gaa gtg gcc atc	24719
Pro Leu Trp Gly Leu	Ala Glu Trp Ala Glu	Arg Glu Val Ala Ile	
1825	1830	1835	
1840	1845	1850	
ctc ctc cgg ctc aag	cac tcc aat ctg gtg	cgc atc cgt ggg cat	24764
Leu Leu Arg Leu Lys	His Ser Asn Leu Val	Arg Ile Arg Gly His	
1855	1860	1865	
gga cag tgg ccg gat	gag gcg cct caa tcc	ttc ttc atc gtc atg	24809
Gly Gln Trp Pro Asp	Glu Ala Pro Gln Ser	Phe Phe Ile Val Met	
1870	1875	1880	
gac tac gtg gaa ggg	cgc cgg ttg gac gta	tgg gcc aga gag gag	24854
Asp Tyr Val Glu Gly	Arg Arg Leu Asp Val	Trp Ala Arg Glu Glu	
1885	1890	1895	
aac ccc tcg gcc cga	gaa gtc gtg ctc aag	gtg cgt ggc gtg gcg	24899
Asn Pro Ser Ala Arg	Glu Val Val Leu Lys	Val Arg Gly Val Ala	
1900	1905	1910	
cgc ggg ctg ggc gcc	gcg cac cgg gcg aag	gtg gtg cat cga gac	24944
Arg Gly Leu Gly Ala	Ala His Arg Ala Lys	Val Val His Arg Asp	
1915	1920	1925	
ctg aag gag agc aac	gtc atc gag cgc acc	tcc gac gga gag gca	24989
Leu Lys Glu Ser Asn	Val Ile Glu Arg Thr	Ser Asp Gly Glu Ala	
1930	1935	1940	
gtg gtg gtg gac ttc	ggc gcg ggc ggg tac	gag agc gcc ccc agc	25034
Val Val Val Asp Phe	Gly Ala Gly Gly Tyr	Glu Ser Ala Pro Ser	
1945	1950	1955	
atc acc ggc ggc gtg	ctg cca ccg ggc acg	ccg gag tac cgc gcg	25079
Ile Thr Gly Gly Val	Leu Pro Pro Gly Thr	Pro Glu Tyr Arg Ala	
1960	1965	1970	
ccc gag gcc tgg cgc	ttc cag cag gag cac	ggg gac gag cgt ggc	25124
Pro Glu Ala Trp Arg	Phe Gln Gln Glu His	Gly Asp Glu Arg Gly	
1975	1980	1985	
cac tcc tac cag ccc	ggc ccc tcg gat gac	ctg tac tcg ctg ggc	25169
His Ser Tyr Gln Pro	Gly Pro Ser Asp Asp	Leu Tyr Ser Leu Gly	
1990	1995	2000	
gtc gtt ctc tat tgg	ctc ctg acg ggc agg	cag ccc ttc ctc ccg	25214
Val Val Leu Tyr Trp	Leu Leu Thr Gly Arg	Gln Pro Phe Leu Pro	
2005	2010	2015	
gac gag gcc gca ggg	gtg gag gcc gtg ctc	aac cgc gcc ccc aaa	25259
Asp Glu Ala Ala Gly	Val Glu Ala Val Leu	Asn Arg Ala Pro Lys	
2020	2025	2030	
cca ccc cat gtg ctc	aat cca cgc gtc ccc	gag gcc ctg agt gcc	25304
Pro Pro His Val Leu	Asn Pro Arg Val Pro	Glu Ala Leu Ser Ala	
2035	2040	2045	
gtg tgc atg cgc ctg	ctg gcc aag gag ccc	gag gag cgc cac ccg	25349
Val Cys Met Arg Leu	Leu Ala Lys Glu Pro	Glu Glu Arg His Pro	
2050	2055	2060	
gac gcc gac gcg ctg	tgc gcg gag ctg gag	gcc ctg ctg gcc cag	25394

Asp Ala Asp Ala Leu	Cys Ala Glu Leu	Glu Ala Leu Leu Ala Gln	
2065	2070	2075	
gcg gac gag tcc tgg	gac gtg aaa ctc tgc	gac gcc tat ggc ccg	25439
Ala Asp Glu Ser Trp	Asp Val Lys Leu Cys	Asp Ala Tyr Gly Pro	
2080	2085	2090	
gac acc gcc acc acg	ctc gcg gcg gtg ccg	cac gcg ggc gaa gac	25484
Asp Thr Ala Thr Thr	Leu Ala Ala Val Pro	His Ala Gly Glu Asp	
2095	2100	2105	
gag ctg gtg caa tgg	ctg aag aag cgc aag	gcc cgg cct cgt cgg	25529
Glu Leu Val Gln Trp	Leu Lys Lys Arg Lys	Ala Arg Pro Arg Arg	
2110	2115	2120	
ggg cca cgt ccg cca	cac ggc gag ggc gca	cac gag ccg gac tca	25574
Gly Pro Arg Pro Pro	His Gly Glu Gly Ala	His Glu Pro Asp Ser	
2125	2130	2135	
aac gcg gtg gcc atc	cag gca gag ctg cca	ccc cca gtt cct gct	25619
Asn Ala Val Ala Ile	Gln Ala Glu Leu Pro	Pro Pro Val Pro Ala	
2140	2145	2150	
cgc ccc ggc ccc atg	ccc tcg ccg ccc ccg	gca caa gcc ggt cac	25664
Arg Pro Gly Pro Met	Pro Ser Pro Pro Pro	Ala Gln Ala Gly His	
2155	2160	2165	
ttc aag gcc ctg cgt	gtg gcg acc tgg atg	ggt ctg ctg gtg gtc	25709
Phe Lys Ala Leu Arg	Val Ala Thr Trp Met	Gly Leu Leu Val Val	
2170	2175	2180	
ttg agc gcg ggc gcg	ctg gtg atc gcg cgc	acc ttg gcg cct tcc	25754
Leu Ser Ala Gly Ala	Leu Val Ile Ala Arg	Thr Leu Ala Pro Ser	
2185	2190	2195	
cct gct cct gcc gca	cgt cac gcc ctg gag	cct cac ccc agt gat	25799
Pro Ala Pro Ala Ala	Arg His Ala Leu Glu	Pro His Pro Ser Asp	
2200	2205	2210	
gcc gtg gac tca cca	cct gca tcc ccc gag	ttc gac ctc gcg ccc	25844
Ala Val Asp Ser Pro	Pro Ala Ser Pro Glu	Phe Asp Leu Ala Pro	
2215	2220	2225	
tgg gcc ccc ggc cag	gaa gtg gcg gcg cca	tgg atg gca ccg gaa	25889
Trp Ala Pro Gly Gln	Glu Val Ala Ala Pro	Trp Met Ala Pro Glu	
2230	2235	2240	
gct cag gaa gcc gca	gcc gcg ctc gat agc	gcg ccc acc cta gcg	25934
Ala Gln Glu Ala Ala	Ala Ala Leu Asp Ser	Ala Pro Thr Leu Ala	
2245	2250	2255	
gcc gtc gcc ctc ccc	gcg acg atc tcc gag	gag aag gct tcc gtg	25979
Ala Val Ala Leu Pro	Ala Thr Ile Ser Glu	Glu Lys Ala Ser Val	
2260	2265	2270	
aag atg aag aag aac	acc ggc gta ttc tcc	gag ccc gag ccg cag	26024
Lys Met Lys Lys Asn	Thr Gly Val Phe Ser	Glu Pro Glu Pro Gln	
2275	2280	2285	
cgc ttc cgg ggc gga	acc gtt ggc aag gcg	ctc ggc ctc ggt gtc	26069
Arg Phe Arg Gly Gly	Thr Val Gly Lys Ala	Leu Gly Leu Gly Val	
2290	2295	2300	

gcg ggc tgc cta gcg	atg gcc tgc cct ggc	ccc cag gtc cgc ccg	26114
Ala Gly Cys Leu Ala	Met Ala Cys Pro Gly	Pro Gln Val Arg Pro	
2305	2310	2315	
acg cct ccg cca gag	gcg tgc ccg cca ggt	gct gtc gag gcc atg	26159
Thr Pro Pro Pro Glu	Ala Cys Pro Pro Gly	Ala Val Glu Ala Met	
2320	2325	2330	
agc agg ctt cga att	gag tat gac gag gat	gtg gta gcc acc ttc	26204
Ser Arg Leu Arg Ile	Glu Tyr Asp Glu Asp	Val Val Ala Thr Phe	
2335	2340	2345	
ttc gcg acc gcc gag	gag ggg gtt agc agg	aac atc acg gtg cgc	26249
Phe Ala Thr Ala Glu	Glu Gly Val Ser Arg	Asn Ile Thr Val Arg	
2350	2355	2360	
gcg ggc ccg aca acg	gtt cac ctg ggg aga	cca ttg gga aat cta	26294
Ala Gly Pro Thr Thr	Val His Leu Gly Arg	Pro Leu Gly Asn Leu	
2365	2370	2375	
ccg gcc ccg acc acg	ctc tcc gga cgc ctc	atc atc gga gag gat	26339
Pro Ala Arg Thr Thr	Leu Ser Gly Arg Leu	Ile Ile Gly Glu Asp	
2380	2385	2390	
cgc gtt tac ggc cga	cta act gaa gca cgc	acg cca aag ggt gac	26384
Arg Val Tyr Gly Arg	Leu Thr Glu Ala Arg	Thr Pro Lys Gly Asp	
2395	2400	2405	
cgc ttc cct gta tgc	atg cag atc att gag	aac gac tct att ggg	26429
Arg Phe Pro Val Cys	Met Gln Ile Ile Glu	Asn Asp Ser Ile Gly	
2410	2415	2420	
ttc aag aag gag atc	aac gac ggc tcc aac	agc gcc tcg gtc tat	26474
Phe Lys Lys Glu Ile	Asn Asp Gly Ser Asn	Ser Ala Ser Val Tyr	
2425	2430	2435	
ccc tat ttt gac gtg	aag gcc gtc cgc cgc	ttc gag tga accccgacca	26523
Pro Tyr Phe Asp Val	Lys Ala Val Arg Arg	Phe Glu	
2440	2445		
ggagcaacat ggggctgaac	caagtatcaa tgcagggttc	cactctgcta cttcaaggtc	26583
tccgcgcgatg ccc atg	tca ccc atc gtt gtt ctc	ctg gtg ata gcg ctt	26632
Met Ser Pro Ile Val	Val Val Leu Leu	Val Ile Ala Leu	
2450	2455		
ttc tcc ggg gca gcc	gcc gcg caa tcg cgt	cca gca gcc acg ggc	26677
Phe Ser Gly Ala Ala	Ala Ala Gln Ser Arg	Pro Ala Ala Thr Gly	
2460	2465	2470	
ttg ggg gta cgc ccg	att gag ttc tcg gcc	gag aac ttc gag gca	26722
Leu Gly Val Arg Arg	Ile Glu Phe Ser Ala	Glu Asn Phe Glu Ala	
2475	2480	2485	
atg acg aca ccg gag	gta cag ata agc ccg	ggg gtt tcc acc acg	26767
Met Thr Thr Pro Glu	Val Gln Ile Ser Pro	Gly Val Ser Thr Thr	
2490	2495	2500	
ttc gag ttc aac tcc	gcc ctc ctt cag gag	aag gtc gcg gtt gaa	26812
Phe Glu Phe Asn Ser	Ala Leu Leu Gln Glu	Lys Val Ala Val Glu	
2505	2510	2515	
ggc gcc gac cgc ttc	tcg ctc gtg gac att	gga cga agc act ctc	26857

Gly 2520	Ala	Asp	Arg	Phe	Ser 2525	Leu	Val	Asp	Ile	Gly 2530	Arg	Ser	Thr	Leu	
cgg	ctg	gtg	ccg	tcg	gag	caa	gtc	ctg	ccc	ggc	gaa	cgc	ctg	cga	26902
Arg 2535	Leu	Val	Pro	Ser	Glu 2540	Gln	Val	Leu	Pro	Gly 2545	Glu	Arg	Leu	Arg	
gtg	acg	gtg	cgg	ttc	ctg	gac	gga	gca	gcg	ccg	gta	ggc	atg	gca	26947
Val 2550	Thr	Val	Arg	Phe	Leu 2555	Asp	Gly	Ala	Ala	Pro	Val	Gly	Met	Ala	
ttc	atc	ctc	gtc	gcg	cac	ccc	gct	cag	gcg	gag	cgc	ttg	gtc	gag	26992
Phe 2565	Ile	Leu	Val	Ala	His 2570	Pro	Ala	Gln	Ala	Glu 2575	Arg	Leu	Val	Glu	
gtg	tgg	cgc	aac	cag	cgc	acg	gtg	gag	tcc	tat	caa	cag	gaa	tcg	27037
Val 2580	Trp	Arg	Asn	Gln	Arg 2585	Thr	Val	Glu	Ser	Tyr 2590	Gln	Gln	Glu	Ser	
aaa	gaa	gcc	cga	gcg	gag	gcc	caa	agg	tgt	cac	gaa	gag	aac	gag	27082
Lys 2595	Glu	Ala	Arg	Ala	Glu 2600	Ala	Gln	Arg	Cys	His 2605	Glu	Glu	Asn	Glu	
cgc	ctg	cgc	gcc	gag	tat	gcg	gga	ccg	ggc	ggc	ctc	gct	ggc	ctc	27127
Arg 2610	Leu	Arg	Ala	Glu	Tyr 2615	Ala	Gly	Pro	Gly	Gly 2620	Leu	Ala	Gly	Leu	
ctt	gcg	aac	cgc	gtg	atc	ggc	aag	tcg	ggc	gtt	tcg	gtc	aag	ccc	27172
Leu 2625	Ala	Asn	Arg	Val	Ile 2630	Gly	Lys	Ser	Gly	Val 2635	Ser	Val	Lys	Pro	
ctc	gac	ttc	gag	aag	gaa	gtc	cgc	cag	cgt	cct	gga	gac	gcg	atc	27217
Leu 2640	Asp	Phe	Glu	Lys	Glu 2645	Val	Arg	Gln	Arg	Pro 2650	Gly	Asp	Ala	Ile	
agg	att	ctt	cgt	gca	tgg	agc	tac	cgc	tcc	gcc	aac	cgc	gtg	tcc	27262
Arg 2655	Ile	Leu	Arg	Ala	Trp 2660	Ser	Tyr	Arg	Ser	Ala 2665	Asn	Arg	Val	Ser	
gtg	ctc	atg	gat	ttg	gac	tac	cca	gaa	gcc	gcg	caa	cgc	tgg	atc	27307
Val 2670	Leu	Met	Asp	Leu	Asp 2675	Tyr	Pro	Glu	Ala	Ala 2680	Gln	Arg	Trp	Ile	
gcg	cag	gga	gca	tcg	ttg	gtg	agc	aag	aca	ggt	gaa	acg	ctg	agc	27352
Ala 2685	Gln	Gly	Ala	Ser	Leu 2690	Val	Ser	Lys	Thr	Gly 2695	Glu	Thr	Leu	Ser	
atg	ctc	cct	gta	tgg	cag	gaa	gcg	ccc	gtt	gcc	aat	gcc	cgg	atc	27397
Met 2700	Leu	Pro	Val	Trp	Gln 2705	Glu	Ala	Pro	Val	Ala 2710	Asn	Ala	Arg	Ile	
cgg	cgc	ctc	gtg	gtg	gaa	gca	gag	gct	gag	ccg	gac	gcg	gcc	cag	27442
Arg 2715	Arg	Leu	Val	Val	Glu 2720	Ala	Glu	Ala	Glu	Pro 2725	Asp	Ala	Ala	Gln	
gac	acg	ttc	acc	ctc	aaa	ctg	tgg	gaa	tcg	aac	ggg	ctg	cgc	acg	27487
Asp 2730	Thr	Phe	Thr	Leu	Lys 2735	Leu	Trp	Glu	Ser	Asn 2740	Gly	Leu	Arg	Thr	
atc	acc	cta	tcc	ggg	gtg	aca	ttc	cca	taa	tggtgagaac	atcctcacta				27537
Ile 2745	Thr	Leu	Ser	Gly	Val 2750	Thr	Phe	Pro							

tcaagctagc acccacatag catgaagtcc gcgcccgtgc gtgtcatggg ccgttcaaac 27597
acctgaaaac caataactac atccttacac catccaaacc aagcccatg tctacagccg 27657
agcaaacgaa gaaaataatc aagaagctgc gagccaccct aaacgaagga agctcagaaa 27717
agattcaagc actcatccag tccatcagac tggagaatcc atctcagtgg ctgaatttga 27777
tcgccctcgc actgggcccc atagatccta ggaatccaac attcttcgaa aaagtatccg 27837
caataaacag attggcagaa gatcttgccg cagaagagca gcacacgcta cgcggatacg 27897
aaagcatcgc aagggccatg cagatcgtat cgcaagagtc cgacaaatca gaattctgga 27957
aatcttcgcg cctcacaatg atgtggcgcc tatttacgta cgttgaatcc caatacatat 28017
ttgttagcaa tacgattcga aagaagctca agcgggcccc accgggctcg ccatttgtca 28077
tggcgagccc ttcaatcgcg gttccaggaa ccgatgggac cgcaacgaca agtgtcgatc 28137
ttgtttttga atctaccgcc gaagccgccg aactaatact ccgcatctcc atgcacagga 28197
tcgacaccca acccccaa at gaaaagtta cccccagac cccttatgac gatcctgact 28257
ttggcaggct cctcgcttgc gcacacgctg ggcacgatt cagcctgcta tggaacaaa 28317
tccagtatac cgactggcga cctgagcatg atccacgagg cttcatctgg acacaaaaa 28377
acaaggcgca atacctcaa agcgaaatcg gaaacattcg cgagcgcta tttttgcagc 28437
aagcaatgct gtctgatgga ctcggtcctg gggcagaact ggagagtaga agagaagtcc 28497
tggaacgcct caagaggagc ttcaatctcc cgaagccagg gaaaacatgg gaccttaaga 28557
ttgatattaa agccctgcga gatctcctcg gtcttgatcg atatgccag gctcgcaaga 28617
tctacatgtc gcgtcgattc ctaacccagc taatgggagc gccagcagat cctgaactgg 28677
agaagtcaca gaaggagttg aagttgaact tcacgctgca tgaactcgcc atggctctgc 28737
tgcaaaggaa tcacgcaaca ttcaagacca agtcagcaat acacaacatt gtttttgcag 28797
tcaccaagca ggcctatgc aaggtaatca gcgaggcgtc gggcatcgaa ttcgaattct 28857
gcttttccca tatcaacaac atcgtcttcg acgcacagag aaaaaaccta gagatctggg 28917
accagccgct cgtaccagta gatgaccata ggggtgttgtt cgttccttcg atcattgcca 28977
attcccatcc tattcgaca atagagcatt tagcttcagc tcgctcaagc tccgcgtttg 29037
actcccgagg agagtattat gagcagatga ttcaatcgag gcttgaggaa ttaacacaag 29097
cagtcgctgt gcggggcatc aaatttcctg cttcagatgg gcggcagatc gaatacgaact 29157
tggtgcttta ttggcaaggg cacctattta tcttcgagct taagtgcctg aagatgacaa 29217
caactcctgt ggatgagttc ggggcgaggg aggagatcga gcgcgccgag caccaattgg 29277
tacggcgctg ttctattgcc gaaactgact ggaaggagtt caggagccaa gcacagggag 29337
tttcaactccc cgagactccg ccagcgcgca gaaacatcca tgccataata ctgaccaaca 29397
tcatgaactt tacgggcccga agaactggag agatcgtgac ggcagacgac tcctgcctgt 29457

ttagatactt	cgacagccct	catgttcccc	tatttgatgc	tgccacgggt	gccgaaatct	29517
caagcattgc	acgtttgcg	cttcgcaacg	aggggccgcc	atctgcttcg	gggctttggc	29577
aatattttatc	aagccctccg	cagttggagt	ggatgcggaa	gcaaattcgg	cctgaatacc	29637
accactgcc	accgctggat	gagggcaact	atcactggaa	gctcatggtc	gcattcgatc	29697
cggccgacgc	ccagacgact	aaatagtgtt	ggtttgggtct	cgttgcgaag	gtgagcattg	29757
acctcggctg	gacgctggat	tcagcgtcta	caacaacgat	attgcagggg	gagggcggagg	29817
ctgcgtagcc	ttcgtttaga	acccgtggag	aacccgaccc	atg acg cta tac gaa		29872
				Met Thr Leu Tyr Glu		2755
ctg ttt	gaa gcc gat gat gac	cgc gat gtc ccc gcc	gac ttt cct			29917
Leu Phe	Glu Ala Asp Asp Asp	Arg Asp Val Pro Ala	Asp Phe Pro			
2760		2765	2770			
tgg ctc	ttg atc gag atc cct	gag gag caa ctc cgc	cat gcg gcg			29962
Trp Leu	Leu Ile Glu Ile Pro	Glu Glu Gln Leu Arg	His Ala Ala			
2775		2780	2785			
ctt ctc	ctg aat gag ggt gag	tgg cga ccc tcg cac	att tct ggc			30007
Leu Leu	Leu Asn Glu Gly Glu	Trp Arg Pro Ser His	Ile Ser Gly			
2790		2795	2800			
atc tgg	tat cgg gtg gat ccc	gag cgc cca gct cag	aag cag cag			30052
Ile Trp	Tyr Arg Val Asp Pro	Glu Arg Pro Ala Gln	Lys Gln Gln			
2805		2810	2815			
cgg cac	gtc cac gtc gcc gcg	aag aag cac atc aag	atc ccc acg			30097
Arg His	Val His Val Ala Ala	Lys Lys His Ile Lys	Ile Pro Thr			
2820		2825	2830			
aag cag	gct tca tgg aac aga	gac acc acc cgc cat	gac cgc aag			30142
Lys Gln	Ala Ser Trp Asn Arg	Asp Thr Thr Arg His	Asp Arg Lys			
2835		2840	2845			
aca ttc	aat gcg aag ctg ggc	tcg caa ggg agc tat	caa gac gtg			30187
Thr Phe	Asn Ala Lys Leu Gly	Ser Gln Gly Ser Tyr	Gln Asp Val			
2850		2855	2860			
gcg aag	acg gcg ctg ggc ctg	ccg ccc gag gct gtc	ctc gaa cac			30232
Ala Lys	Thr Ala Leu Gly Leu	Pro Pro Glu Ala Val	Leu Glu His			
2865		2870	2875			
gtt gtg	gaa aat ccg gcc gaa	caa cag atg ctg ctc	acc gaa tcc			30277
Val Val	Glu Asn Pro Ala Glu	Gln Gln Met Leu Leu	Thr Glu Ser			
2880		2885	2890			
act gat	ccc cag ttg aaa act	gag ttc tac agg tgg	ttg gag gcc			30322
Thr Asp	Pro Gln Leu Lys Thr	Glu Phe Tyr Arg Trp	Leu Glu Ala			
2895		2900	2905			
cag ccc	agg cgg tct aac ttc	gaa gcg ctc gtc gag	gac gcg gtt			30367
Gln Pro	Arg Arg Ser Asn Phe	Glu Ala Leu Val Glu	Asp Ala Val			
2910		2915	2920			
agg ctc	aat ctg ccg ctc gac	ctg gaa tcg taa ggcagtggtt				30410
Arg Leu	Asn Leu Pro Leu Asp	Leu Glu Ser				

2925

2930

```

aatgcggtg atacgagcgg tgatacgagc gcctaacaaa actggcatag ggaccccggt 30470
tctcaaggac gcagctgcaa ccaaggcgat tttgtaggca ttctttgggg cccgagacag 30530
cgcgcgggat ctccggcacc cagttcacgg acgtcacctg ccccggtgca gccaatggcc 30590
cctgaccctt gccgtcccca ccctcgtagc caagctgtac ccaagatgaa cggacgcggg 30650
tggacctcca tggacgctcc acccaccctg ctccagaggg aaaccggccg accctcaatc 30710
cctccggggca cctcgcggtat cgggggagcg gggcttcaaa accggtgagg ggcgctgaga 30770
ggcgctccctg gtgagttcga ttctcatggg gcgacgtgaa cccggccgtg ctccatgcgc 30830
tccacaacgc ggcggaagat gaggtcgata atggcgaggg actgagcctt caggagtggc 30890
cggaacagac cagggccacg tcagaagcga acgacctccc ttctgtaacc gaataacgaa 30950
acagaggcga ggatcacaac ttctgaaatc cgttgacgcc cgtgggcttc atttcaggaa 31010
tcctgggaaa tgctctcct tgctcttctt gggcttctcc atggtagcgc ttcgcgcacc 31070
caggctctgg tactatccat catatggagc gctccctcgg aacgacagag caggtcttct 31130
ataaggtacg gcagaagcgc cccttcaaca tcgtcattac gggcaagctg cgaggacgcg 31190
tggaggagca ggctgtgcgc caggggttg gtg gcc ctg caa cgg cga cac ccg 31243
                               Val Ala Leu Gln Arg Arg His Pro
                               2935                               2940

ctg ctc cga gcg aag atc gtc gcg ggt act ccg cca cgc ttc acc 31288
Leu Leu Arg Ala Lys Ile Val Ala Gly Thr Pro Pro Arg Phe Thr
                2945                               2950                               2955

tcg caa ggt gtg ccc ccc att gcg ctg gaa gtg gtg gaa cgc aag 31333
Ser Gln Gly Val Pro Pro Ile Ala Leu Glu Val Val Glu Arg Lys
                2960                               2965                               2970

gca gag gag tcc tgg caa cag cat gtc gag gca gag ctt gac agg 31378
Ala Glu Glu Ser Trp Gln Gln His Val Glu Ala Glu Leu Asp Arg
                2975                               2980                               2985

cct ttc aac tcg gac cct ggg ccg ctc gtc cgc ttc atc ctc gtg 31423
Pro Phe Asn Ser Asp Pro Gly Pro Leu Val Arg Phe Ile Leu Val
                2990                               2995                               3000

cgg ggc gag cac cgc tcc gag ctg ctg tgc agc tat gat cat ctg 31468
Arg Gly Glu His Arg Ser Glu Leu Leu Cys Ser Tyr Asp His Leu
                3005                               3010                               3015

atc gga gac gcg cac tcc ggc atc ttc gcc ctg cgc gat ctg ctg 31513
Ile Gly Asp Ala His Ser Gly Ile Phe Ala Leu Arg Asp Leu Leu
                3020                               3025                               3030

cag gtc atg gcc tcc cag gac cag cac ctg ccc gag ctc gcc cca 31558
Gln Val Met Ala Ser Gln Asp Gln His Leu Pro Glu Leu Ala Pro
                3035                               3040                               3045

agg ccc gcc tat gag gag ctg att ggg ccg atg gta ccg ggc acc 31603
Arg Pro Ala Tyr Glu Glu Leu Ile Gly Pro Met Val Pro Gly Thr
                3050                               3055                               3060

```

cga ctg ctt ggc Arg Leu Leu Gly 3065	gcc gcc gtg cgc cgg Ala Ala Val Arg Arg 3070	ggc tcc tcg gcg ctg ctg Gly Ser Ser Ala Leu Leu 3075	31648
cgc ttg tct ccg Arg Leu Ser Pro 3080	acc ctc aat acg ctc Thr Leu Asn Thr Leu 3085	aca gag cgg ctc gtc gca Thr Glu Arg Leu Val Ala 3090	31693
ccg cgg ggc atg Pro Arg Gly Met 3095	gct cgc agc agc gcc Ala Arg Ser Ser Ala 3100	gag gga cag gcg ctc tac Glu Gly Gln Ala Leu Tyr 3105	31738
agc cac cga att Ser His Arg Ile 3110	ctg gag ccc gag cag Leu Glu Pro Glu Gln 3115	ctc gcg cgg ctg ctt gcg Leu Ala Arg Leu Leu Ala 3120	31783
cgc tgt cga gag Arg Cys Arg Glu 3125	caa ggg tcc tcc gtc Gln Gly Ser Ser Val 3130	cac gca gcg ctc ggg gcc His Ala Ala Leu Gly Ala 3135	31828
gcg ctg ctg ata Ala Leu Leu Ile 3140	gct cgg gcc gag tcg Ala Arg Ala Glu Ser 3145	cag ggt gcc aag aag cg Gln Gly Ala Lys Lys Arg 3150	31873
gtg acg ctc acc Val Thr Leu Thr 3155	ctc acc tct gcg ttg Leu Thr Ser Ala Leu 3160	gat gct cgg gag cgc ttt Asp Ala Arg Glu Arg Phe 3165	31918
ggc gtg ggt gag Gly Val Gly Glu 3170	gac ttc ggc ctg ttc Asp Phe Gly Leu Phe 3175	acc acg ggg aag acg gac Thr Thr Gly Lys Thr Asp 3180	31963
ttc ttc cgc gcg Phe Phe Arg Ala 3185	cgc ggc agc acc cca Arg Gly Ser Thr Pro 3190	ttc tgg gag ctg gcg cac Phe Trp Glu Leu Ala His 3195	32008
aga ctc cga tcg Arg Leu Arg Ser 3200	cct ctg cag gca gca Pro Leu Gln Ala Ala 3205	cgg aag aag cgg agc cac Arg Lys Lys Arg Ser His 3210	32053
ttg agg ctg ttc Leu Arg Leu Phe 3215	cgc ctg atc ctg gag Arg Leu Ile Leu Glu 3220	ata tcc gcg ctc acc gtg Ile Ser Ala Leu Thr Val 3225	32098
gac atc tcc tcc Asp Ile Ser Ser 3230	cag ccc tgg atg gaa Gln Pro Trp Met Glu 3235	cgc gcg aca cgg ctg ggc Arg Ala Thr Arg Leu Gly 3240	32143
ctg cac tcc atg Leu His Ser Met 3245	ctc gcc ctg agc aac Leu Ala Leu Ser Asn 3250	gtg ggg cgt gtg gag att Val Gly Arg Val Glu Ile 3255	32188
gcc gcc cgc tat Ala Ala Arg Tyr 3260	ggg aac ctg acg ctc Gly Asn Leu Thr Leu 3265	gag ggc ctt ggc ttc acg Glu Gly Leu Gly Phe Thr 3270	32233
ggg agc acc gcg Gly Thr Thr Ala 3275	gcc cag ttc gac gtg Ala Gln Phe Asp Val 3280	gtg ctg acg gcg atc acc Val Leu Thr Ala Ile Thr 3285	32278
ttt gcc cag cgg Phe Ala Gln Arg 3290	ttg gag gcc agc ttt Leu Glu Ala Ser Phe 3295	ctc ttc aac acc gcg cac Leu Phe Asn Thr Ala His 3300	32323

3290

3295

3300

atg cag cgc gag cag gtg gag cag ctc gcc tcg cgc acc tgg gag 32368
Met Gln Arg Glu Gln Val Glu Gln Leu Ala Ser Arg Thr Trp Glu
3305 3310 3315

ctg ctg gcc gag gcg acg cgc tga gcggctactc cggctggatg ttgtacatcc 32422
Leu Leu Ala Glu Ala Thr Arg
3320

gcttgacctc ttccagcggc agctccatga ccttccaggg atcccagtct ccgaagaggt 32482
ccacgctcat cgcgctgcct cgctccaccg cacgccgcag ctggggaatg ggaagcgcct 32542
cggcgagcag gcgaacgggc accaaggact tgtactgggc cggatcatag ttgcgggtgc 32602
gggggcccga catgaagccc aggctgaagg ccgtggtcag catttcctcc cagatggaga 32662
tgccatagcc ggagagcagg tggacgaagt catgcacggc cgagtgcacg ggcagggcgt 32722
gagggtgacc cggcagcggc acgccatggg agtggcagta cttccagaac tcgcgggccaa 32782
aggaaccctc gggcaggtgc tcgagctgct gaaagcgttg cttgaggtct ggcgcctcat 32842
tgcgctcgcg cagcatgtca tagagcttcc agggcagctc gcccttcggc agccccaaca 32902
tgcggcagaa gtactcaata taggcaggca cgaggaaggc ggtcagtcca acggacagcc 32962
acagcttgga gcgcgcgtcc gcggccaggg aacgctcgcg cgggtccaggg gctcgcctgg 33022
acttccgcag aaatacatcc gtgagtattc tcattcgaac accttggtc gttatacggc 33082
tggggcagct ttcaggcagg tcgctgtctc ctggagaaag ccatacggat ccttggggtc 33142
ctcggccatg gactccacct ggagcaggcg gcctaccccc agccgctcgg cgtgcgccag 33202
gatgacttcc agcaccacct tgtcctccaa ggaaaaccgc gtggagtcga agacgggtgag 33262
gccatcgcgg tgagagcgcc actgctcggg agagcgggtc agttcaaaga ggctggggcc 33322
aatctctcca ggctcgagct gctggcactc cccctccacc acggcctggg ggagatagtc 33382
cgggcatacc agggcgcgac gcagcatctc caggggcagc tctgtcttgc cgggcagatc 33442
tgagcccacg gcgttgatat ggaggtgggc cttgatcaga cgcccgaca ccaccggccc 33502
cgcgcccacg gcgacggagg tggccgtgca gaggatgtct gcctcctgct ccacgcgcgt 33562
caggggggacc gccaccacct ccaggcccag gaaggccaca cgccgggcaa agcttcgcat 33622
ggcttcggga gcaatgtcgt gcaccagcac ccgtcccaga ggaaagacgc gcgagagcgc 33682
atgtgcctgg gtgacggcct gagcgccaca ccctaccagg cccaataccc ggctgtcctc 33742
ccgcgcgagc agccgactcg ccacggcgga ggccgccccg gtgcgtagcg cggtaggaaa 33802
ggtgccatca caaacggccc gcaggtgccc ggtcgtcaca tcatacagac tgagggagga 33862
aagaatggtg ggcaagccca gtcctccgg gctgcggggg ttgtagccca ccagcttgat 33922
cacggcgtgg cgccccacct gcatgacggg catccactcc aacacgccgg ggtgggggtg 33982
ctgatactgg aagccctcgc gcttgccgag ctccggtgtg gcggggtcga agtgctccag 34042

ctcgtactcg	aggggtggaga	tgacctcatc	catcagcgcg	tcgagccccc	ttgcacggac	34102
gatctgggcgc	agttcctgct	cacccaggat	ccttgtcatc	cgcattggctg	accccgcgcc	34162
tttcagctgc	atgacaggcc	tcaactgcatt	cacggctcca	gttcaggggcc	gggcaaaatc	34222
tatcttcaac	atccccctgt	ccaggaccgc	gtccacggca	tccagctgcg	ccaccgagcg	34282
cggcagcatg	acgtggcggc	ggaagccgcc	cacgcggatg	atgaggtcat	tgccctggcg	34342
atccaggaca	atgtctccgc	ggtcggcacc	tggcatgagg	agctgcagcc	gatagcggcc	34402
ctcgtgctgc	ttggcgaact	gatagggctg	cccggagatg	tagcgctcgg	tgggggtccga	34462
gtccccgtag	agccgctgcg	ccagatcctc	caaccgctcc	agccccacca	cctcgtgctc	34522
cagcagcggg	agcctggcca	ccggcatggg	ggagaagtac	tcttggatgt	gatccaggta	34582
ggcagattga	gacttgtgcc	actgagcaaa	gtattccgtg	tccggcagca	gacgattgac	34642
cacgacctgg	tccaccgtca	tgccatacat	gttgaagtag	aggtaggccc	gctgtgtctc	34702
gcggatgacc	atTTTTTccg	cgtctggagac	caatcgcacc	gtggtgcgat	tgccatccgt	34762
gagcaaccct	tcaatgccct	tgacgcgatc	gaagagctgc	tcgagggagg	caaagtagct	34822
gtcctccggc	agatcatagc	tcgtgagccg	actggccagg	ggccgggcca	gcttgacgag	34882
cgtgcgatcc	acattgaagc	gccggcggat	ataccactgc	agcacgctgg	tgatgttgac	34942
gaagcgcagc	gactcgcccc	tgggcggaca	gtccacgatg	atcacgtcat	agcgcacctg	35002
ctgcacgtac	tggttcaggt	agatgagcga	gatgacgtcc	tccgtgcccc	ggagaatggc	35062
gacctcctcg	gccaccatgt	tggatacccc	ggtggaggtc	agcagcacgg	acatgtagtt	35122
atagacttca	ctccactgcc	gctggagttc	gtgctggatg	tcgatctctt	gcagttcaag	35182
gttggggggcg	accttctggg	gcagtccttc	gttgaagtcg	aagagcttgc	gatcgagatc	35242
gaagctgtcc	gagaggctgt	gggcaatgtc	gaaggacatc	accaaggtgc	ggtagccgcg	35302
cttggcagcg	gccaccgcgg	tggccgcgga	tgccgtcgtc	ttgcctaccc	cacccttgcc	35362
agagaaaaga	agaattcttg	tcattggatgc	ccgcgtgcc	tgccggcaac	gggccgccat	35422
ggcgctctcc	ccctccatgg	gtcacctgtt	gcgtctgtgg	aactcatcac	tatcagcgac	35482
atgccagaag	gagcaagctg	gacggtggct	tgacttgggt	atg acg	ctg gac agt	35536
				Met Thr	Leu Asp Ser	
				3325		
agg tat	gcc aat aga ggt ggg	ggg ctt tat gaa atc	gat gga ttg			35581
Arg Tyr	Ala Asn Arg Gly Gly	Gly Leu Tyr Glu Ile	Asp Gly Leu			
	3330	3335	3340			
gct atc	ggg cac agc ccc ccc	cct tgt tca att cca	gga ggg gtg			35626
Ala Ile	Gly His Ser Pro Pro	Pro Cys Ser Ile Pro	Gly Gly Val			
	3345	3350	3355			
caa ttg	acagcc aca gcg act	gcc gag cac gcg ctg	aaa gcc ctc			35671
Gln Leu	Thr Ala Thr Ala Thr	Ala Glu His Ala Leu	Lys Ala Leu			

3360	3365	3370	
cgg gcc aag gga ctc gcc acg tcg gac gag gcg ctg ccc gcg gtg Arg Ala Lys Gly Leu Ala Thr Ser Asp Glu Ala Leu Pro Ala Val 3375 3380 3385			35716
cgg cac gcc agc ggc gaa aag gct ccc ctc tca tcc tca cag cgg Arg His Ala Ser Gly Glu Lys Ala Pro Leu Ser Ser Ser Gln Arg 3390 3395 3400			35761
cgc atg tgg ttc ctg gag cag ctc gag cca ggc aac gcg gcc cag Arg Met Trp Phe Leu Glu Gln Leu Glu Pro Gly Asn Ala Ala Gln 3405 3410 3415			35806
cat ttg ttg caa agt cat cac ata cag ggg cct ttg cag gta gag His Leu Leu Gln Ser His His Ile Gln Gly Pro Leu Gln Val Glu 3420 3425 3430			35851
ggg ttg cgg cgg gcg ttg ggg ctc atg gtc aag cgc cat gag gcg Gly Leu Arg Arg Ala Leu Gly Leu Met Val Lys Arg His Glu Ala 3435 3440 3445			35896
ctc cgg ctg gtg gtt gcc att ggc gcg gcg ggc cct gag cag cgt Leu Arg Leu Val Val Ala Ile Gly Ala Ala Gly Pro Glu Gln Arg 3450 3455 3460			35941
gtg cgg ccc gag tgg tgg ccc gag ctg cct ctc agc gat ctg agc Val Arg Pro Glu Trp Trp Pro Glu Leu Pro Leu Ser Asp Leu Ser 3465 3470 3475			35986
ggg gtc gcc cag gcg gac agg gct cgg gcc ctg gcg gag ctg gcg Gly Val Ala Gln Ala Asp Arg Ala Arg Ala Leu Ala Glu Leu Ala 3480 3485 3490			36031
cgc gtg gag gct gcg gcg ccc ttc aac ctg cag caa ggc cca ctg Arg Val Glu Ala Ala Ala Pro Phe Asn Leu Gln Gln Gly Pro Leu 3495 3500 3505			36076
ttc cga gtc cgg ctg gcg cga ctg gcg gaa acg gag cat gtg ctg Phe Arg Val Arg Leu Ala Arg Leu Ala Glu Thr Glu His Val Leu 3510 3515 3520			36121
ctg gtc acc ctg cat cac ctg atc tcg gat gga gcc tgg agc tgc Leu Val Thr Leu His His Leu Ile Ser Asp Gly Ala Trp Ser Cys 3525 3530 3535			36166
gag gtg ctc atc aaa gag ctg gcg atg ctc tac ggc cag cac gtg Glu Val Leu Ile Lys Glu Leu Ala Met Leu Tyr Gly Gln His Val 3540 3545 3550			36211
gaa ggc tcg agc ccg gag ctg gcg cca ctg ccc gtc caa tac agg Glu Gly Ser Ser Pro Glu Leu Ala Pro Leu Pro Val Gln Tyr Arg 3555 3560 3565			36256
gac tac gct ggc tgg gag gca agc ttc gcg ccc gcg gga gag gcc Asp Tyr Ala Gly Trp Glu Ala Ser Phe Ala Pro Ala Gly Glu Ala 3570 3575 3580			36301
ttg gag gcc tgg tgg cgg cag cgg ctc gcc ggg gtg ccc acg gtg Leu Glu Ala Trp Trp Arg Gln Arg Leu Ala Gly Val Pro Thr Val 3585 3590 3595			36346
ctg gag ctg ccc acg gag gga gcg cgt ccc ctc aga cag acg tac			36391

Leu	Glu	Leu	Pro	Thr	Glu	Gly	Ala	Arg	Pro	Leu	Arg	Gln	Thr	Tyr	
3600						3605					3610				
cgg	gca	ggg	cgg	gtg	gcc	atc	acc	gtg	cag	ccg	cgc	ctg	cgc	cag	36436
Arg	Ala	Gly	Arg	Val	Ala	Ile	Thr	Val	Gln	Pro	Arg	Leu	Arg	Gln	
3615						3620					3625				
gcg	ctg	gag	gac	ctg	gcg	cac	aag	gag	ggc	gtg	tca	ctc	ttt	gcc	36481
Ala	Leu	Glu	Asp	Leu	Ala	His	Lys	Glu	Gly	Val	Ser	Leu	Phe	Ala	
3630						3635					3640				
ctg	ctg	ctc	acc	gcc	ttc	acc	acg	ctg	cta	cac	cgc	tac	tct	cgc	36526
Leu	Leu	Leu	Thr	Ala	Phe	Thr	Thr	Leu	Leu	His	Arg	Tyr	Ser	Arg	
3645						3650					3655				
cag	gag	gaa	ttg	gtg	gtg	ggg	tgg	ccc	gcc	ccg	cag	cgc	ccc	cgg	36571
Gln	Glu	Glu	Leu	Val	Val	Gly	Trp	Pro	Ala	Pro	Gln	Arg	Pro	Arg	
3660						3665					3670				
ccc	gag	ctg	cat	ggc	ctc	att	gga	tac	ttc	ggc	acc	ccg	gtg	gca	36616
Pro	Glu	Leu	His	Gly	Leu	Ile	Gly	Tyr	Phe	Gly	Thr	Pro	Val	Ala	
3675						3680					3685				
ttg	cgc	tcg	cgg	ctg	gag	cct	cac	acc	cgg	gtg	cgg	gag	gcg	ctg	36661
Leu	Arg	Ser	Arg	Leu	Glu	Pro	His	Thr	Arg	Val	Arg	Glu	Ala	Leu	
3690						3695					3700				
cgg	cag	ctc	gat	cag	gag	gtg	cgt	gag	gcc	aac	gcc	cat	gcg	gcg	36706
Arg	Gln	Leu	Asp	Gln	Glu	Val	Arg	Glu	Ala	Asn	Ala	His	Ala	Ala	
3705						3710					3715				
ctc	ccc	ttc	gag	cgg	ctg	gtg	agc	ctg	cta	gac	att	gcc	cgc	agc	36751
Leu	Pro	Phe	Glu	Arg	Leu	Val	Ser	Leu	Leu	Asp	Ile	Ala	Arg	Ser	
3720						3725					3730				
ccc	agc	cgc	cac	ccg	ctg	ttc	cag	gtg	ctg	ttc	gac	ctg	ctg	ccg	36796
Pro	Ser	Arg	His	Pro	Leu	Phe	Gln	Val	Leu	Phe	Asp	Leu	Leu	Pro	
3735						3740					3745				
gag	cag	ccc	cag	cct	acc	gct	ggc	ggc	tgc	gcc	ttc	cgg	ccc	tgg	36841
Glu	Gln	Pro	Gln	Pro	Thr	Ala	Gly	Gly	Cys	Ala	Phe	Arg	Pro	Trp	
3750						3755					3760				
gag	tcc	ttc	acc	ggc	ctg	gtg	gcg	tat	gat	ctc	act	ctg	ctg	ctg	36886
Glu	Ser	Phe	Thr	Gly	Leu	Val	Ala	Tyr	Asp	Leu	Thr	Leu	Leu	Leu	
3765						3770					3775				
gag	ccg	cgc	ggt	gag	ggg	ctg	gag	ggg	gcg	ctg	gac	tac	agc	gcg	36931
Glu	Pro	Arg	Gly	Glu	Gly	Leu	Glu	Gly	Ala	Leu	Asp	Tyr	Ser	Ala	
3780						3785					3790				
gac	ctg	ttc	agc	gag	gcg	cgc	atg	cag	cgc	gcc	gcc	act	cag	tac	36976
Asp	Leu	Phe	Ser	Glu	Ala	Arg	Met	Gln	Arg	Ala	Ala	Thr	Gln	Tyr	
3795						3800					3805				
ctg	cac	ctg	ctg	gag	cag	ctc	gtg	gag	cgg	cct	cac	gag	agg	ctc	37021
Leu	His	Leu	Leu	Glu	Gln	Leu	Val	Glu	Arg	Pro	His	Glu	Arg	Leu	
3810						3815					3820				
ccg	ccg	ctg	gcg	ctc	ctc	acg	cct	ggt	gag	cac	gag	gct	ttg	ctg	37066
Ser	Arg	Leu	Ala	Leu	Leu	Thr	Pro	Gly	Glu	His	Glu	Ala	Leu	Leu	
3825						3830					3835				

gcg ggg Ala Gly 3840	cac ggc ctg ggc ggg His Gly Leu Gly Gly 3845	ccc gtg gag ggc gcg Pro Val Glu Gly Ala 3850	cag gtg ctg Gln Val Leu	37111
gcc cac Ala His 3855	cac cgc ttc gag cac His Arg Phe Glu His 3860	cag gtg cgg ctc acc Gln Val Arg Leu Thr 3865	ccc cat cac Pro His His	37156
ccc gcg Pro Ala 3870	ttg tgc ttc ggc ccg Leu Cys Phe Gly Pro 3875	cag gtg ctc tcc tac Gln Val Leu Ser Tyr 3880	gag cag ctc Glu Gln Leu	37201
aac cgc Asn Arg 3885	cgc gcc aac ccg ctc Arg Ala Asn Pro Leu 3890	gcc cac cgg ctg cgg Ala His Arg Leu Arg 3895	cgc ctg ggc Arg Leu Gly	37246
gcg ggc Ala Gly 3900	ccg gac acc ctg gta Pro Asp Thr Leu Val 3905	ggg ctg tgc gtg gag Gly Leu Cys Val Glu 3910	cgc tcc ctg Arg Ser Leu	37291
gag ctg Glu Leu 3915	ccg gtg gca ctg ctg Pro Val Ala Leu Leu 3920	gcc ata tgg aag gcc Ala Ile Trp Lys Ala 3925	ggg gcc ggc Gly Ala Gly	37336
ttc ctt Phe Leu 3930	ccg ctg gac gtc aat Pro Leu Asp Val Asn 3935	cag ccc cgc gag cgg Gln Pro Arg Glu Arg 3940	ctc gcc ttc Leu Ala Phe	37381
ctc ctg Leu Leu 3945	gga gat gcg agc tgc Gly Asp Ala Ser Cys 3950	cgc atc ctc ctc acc Arg Ile Leu Leu Thr 3955	cag gag cac Gln Glu His	37426
ctg ctc Leu Leu 3960	cag cgc ctg ccc ccc Gln Arg Leu Pro Pro 3965	acg aac gcc gcg ctc Thr Asn Ala Ala Leu 3970	ctg tgt ctg Leu Cys Leu	37471
gag agg Glu Arg 3975	gag gcg gag gca ctg Glu Ala Glu Ala Leu 3980	gag cgc gag ccc cag Glu Arg Glu Pro Gln 3985	gag gat gcg Glu Asp Ala	37516
cca cac Pro His 3990	gag gcc ggg ctc gac Glu Ala Gly Leu Asp 3995	aac ctg gcc tac gtc Asn Leu Ala Tyr Val 4000	atc cac acc Ile His Thr	37561
tct ggc Ser Gly 4005	tcc acg ggc acg ccc Ser Thr Gly Thr Pro 4010	aag ggc atc gcg atg Lys Gly Ile Ala Met 4015	gtg cat cgc Val His Arg	37606
tgt ctg Cys Leu 4020	gcc aac ctg gtg gcc Ala Asn Leu Val Ala 4025	tgg cag ctc acc cat Trp Gln Leu Thr His 4030	gag cgg ctg Glu Arg Leu	37651
ggt ggc Gly Gly 4035	ccc tca cgc acg ctc Pro Ser Arg Thr Leu 4040	cag ttc gcc tcg ctc Gln Phe Ala Ser Leu 4045	aac ttc gac Asn Phe Asp	37696
atc tgc Ile Cys 4050	tac cag gaa ctc ttc Tyr Gln Glu Leu Phe 4055	acc acc tgg gcc gct Thr Thr Trp Ala Ala 4060	ggg ggc acc Gly Gly Thr	37741
gtg gtc Val Val 4065	atg gtg acg gaa gag Met Val Thr Glu Glu 4070	gtg cgc agg gat cct Val Arg Arg Asp Pro 4075	gcc cgg ctg Ala Arg Leu	37786

ctg gaa Leu Glu 4080	gtg ctg gag cag gag Val Leu Glu Gln Glu 4085	cag gtg agc cgg ctg Gln Val Ser Arg Leu 4090	tac ctg ccc Tyr Leu Pro	37831
ttc atc Phe Ile 4095	gcc ctg cag cag ctc Ala Leu Gln Gln Leu 4100	gcg cgg gtg gcg gat Ala Arg Val Ala Asp 4105	gag cgc ggc Glu Arg Gly	37876
gcg gcc Ala Ala 4110	cct cgc cac ctg cgc Pro Arg His Leu Arg 4115	cag ctc att acc gcc Gln Leu Ile Thr Ala 4120	ggg gag caa Gly Glu Gln	37921
ctc cag Leu Gln 4125	gcg acc ccc gag ttg Ala Thr Pro Glu Leu 4130	cag cgg ctg ctc tcg Gln Arg Leu Leu Ser 4135	cgg atg ccg Arg Met Pro	37966
gag tgc Glu Cys 4140	acc ctc cac aac cag Thr Leu His Asn Gln 4145	tac ggc ccc tcc gag Tyr Gly Pro Ser Glu 4150	tgc cac gtg Cys His Val	38011
gtg acg Val Thr 4155	agc cat gac ctg aca Ser His Asp Leu Thr 4160	cgc gag ccg agc cgt Arg Glu Pro Ser Arg 4165	tgg cca cgg Trp Pro Arg	38056
ctg ccg Leu Pro 4170	ccc gta ggc cgg cca Pro Val Gly Arg Pro 4175	ctg gcc cac ctc cgg Leu Ala His Leu Arg 4180	gtg ctg ttg Val Leu Leu	38101
ctg gac Leu Asp 4185	ggg gag cag cag ctc Gly Glu Gln Gln Leu 4190	gtg cct ccg ggt gtg Val Pro Pro Gly Val 4195	gcc gga gag Ala Gly Glu	38146
gtg ttc Val Phe 4200	ctg ggc ggc ccc gcc Leu Gly Gly Pro Ala 4205	ctg gcg cgc ggc tac Leu Ala Arg Gly Tyr 4210	ctg gga cgt Leu Gly Arg	38191
ccg gag Pro Glu 4215	cag acg gct gat cgc Gln Thr Ala Asp Arg 4220	ttc gtg ccc gat cca Phe Val Pro Asp Pro 4225	ttc tcc cgc Phe Ser Arg	38236
gag cct Glu Pro 4230	gga gcc cgg ctc tac Gly Ala Arg Leu Tyr 4235	cgc acc ggg gac ctg Arg Thr Gly Asp Leu 4240	gcg cgc ctc Ala Arg Leu	38281
cga gag Arg Glu 4245	gac gga gcg ctg gag Asp Gly Ala Leu Glu 4250	ttc ctg caa cgg atg Phe Leu Gln Arg Met 4255	gac gcg cag Asp Ala Gln	38326
gtg aag Val Lys 4260	atc cgg ggc tac cgc Ile Arg Gly Tyr Arg 4265	atc gag cct ggg gag Ile Glu Pro Gly Glu 4270	atc gag gtg Ile Glu Val	38371
gtg ctg Val Leu 4275	tgc gag cac ccc gcc Cys Glu His Pro Ala 4280	gtg cac caa gcg cac Val His Gln Ala His 4285	gtg cgg ccc Val Arg Pro	38416
tac gtg Tyr Val 4290	gac agc gcc gga gag Asp Ser Ala Gly Glu 4295	cgg cgc ctg gtg gcc Arg Arg Leu Val Ala 4300	tac gtg gcc Tyr Val Ala	38461
gca cgg Ala Arg	ctc gag gac acg gac Leu Glu Asp Thr Asp	ggg gcg gag aca gag Gly Ala Glu Thr Glu	cac gtg gag His Val Glu	38506

4305	4310	4315	
cgc tgg cgc gcg gtg tgg gat Arg Trp Arg Ala Val Trp Asp 4320 4325	gag acc tat ggc ggc cca agc ggc Glu Thr Tyr Gly Gly Pro Ser Gly 4330	38551	
acg gcc ttt gac ctg gcc ggg Thr Ala Phe Asp Leu Ala Gly 4335 4340	tgg aat gac agc gtg cgc ggc gag Trp Asn Asp Ser Val Arg Gly Glu 4345	38596	
ccg ctg cct ccg gag cag atg Pro Leu Pro Pro Glu Gln Met 4350 4355	cga gag tgg gtg gag acg aca gtg Arg Glu Trp Val Glu Thr Thr Val 4360	38641	
gag cgg ctc atg gag ctc gtt Glu Arg Leu Met Glu Leu Val 4365 4370	ccc cgg cgg gtg ctg gag ctg ggc Pro Arg Arg Val Leu Glu Leu Gly 4375	38686	
tgt ggc tcc ggg ctc ctg ctc Cys Gly Ser Gly Leu Leu Leu 4380 4385	cgc cgg ctc gcc ccc cgc tgc gag Arg Arg Leu Ala Pro Arg Cys Glu 4390	38731	
tcc tac tgg ggc acg gag ctc Ser Tyr Trp Gly Thr Glu Leu 4395 4400	tcc ccg gta gcc gtc gag cgg ctg Ser Pro Val Ala Val Glu Arg Leu 4405	38776	
agg gag caa ctc cag acg ggc Arg Glu Gln Leu Gln Thr Gly 4410 4415	ggc tcc ccg ctc gcg cag cgc gtg Gly Ser Pro Leu Ala Gln Arg Val 4420	38821	
cgg ctc atg gcc cag ccg gcg Arg Leu Met Ala Gln Pro Ala 4425 4430	gat gac ttc tcc ggg ctt ccg gaa Asp Asp Phe Ser Gly Leu Pro Glu 4435	38866	
gcc ggc ttc gac acc gtc atc Ala Gly Phe Asp Thr Val Ile 4440 4445	ctc aac tcg gtg acg cag ctc ttc Leu Asn Ser Val Thr Gln Leu Phe 4450	38911	
ccc agc gtg gac tac ctg ctg Pro Ser Val Asp Tyr Leu Leu 4455 4460	cgg gtc gtg gag ggc gcg ctg cgc Arg Val Val Glu Gly Ala Leu Arg 4465	38956	
gtg ctc caa ccg ggc gga acc Val Leu Gln Pro Gly Gly Thr 4470 4475	ctg ttc att ggc gac gtg cag aat Leu Phe Ile Gly Asp Val Gln Asn 4480	39001	
ctg cgg ctc ttc gag ctg ttc Leu Arg Leu Phe Glu Leu Phe 4485 4490	cat gcc tcg gtg gcc ctg gag cag His Ala Ser Val Ala Leu Glu Gln 4495	39046	
gcg tca gcg gac ctg gag gcg Ala Ser Ala Asp Leu Glu Ala 4500 4505	cca gcg ctg ctc gcc cgc acc cgc Pro Ala Leu Leu Ala Arg Thr Arg 4510	39091	
cag cgc atg ctg ctg gac gaa Gln Arg Met Leu Leu Asp Glu 4515 4520	cgg ttg tac gtg gac ccg gac ttc Arg Leu Tyr Val Asp Pro Asp Phe 4525	39136	
ttc gcc gcg ctc gcc acg cac Phe Ala Ala Leu Ala Thr His 4530 4535	ttt ccc cag ctg ggc gcg gtc cgg Phe Pro Gln Leu Gly Ala Val Arg 4540	39181	
ctg cac ctc aag cgc ggg agc ggg agg aac gag atg aac cgc ttc		39226	

Leu His 4545	Leu Lys Arg Gly 4550	Ser Gly Arg Asn Glu Met 4555	Asn Arg Phe	
cgc tac Arg Tyr 4560	gac gtg gag ctg cag Asp Val Glu Leu 4565	ctc gcc ccc gta gcg Leu Ala Pro Val Ala 4570	aaa gca ggc Lys Ala Gly	39271
ccc gcg Pro Ala 4575	caa gag ctg cca gag Gln Glu Leu Pro 4580	ctg gac tgg agg cat Leu Asp Trp Arg His 4585	gag ggg ctc Glu Gly Leu	39316
ggc ctg Gly Leu 4590	gaa agg ctg gag cgc Glu Arg Leu Glu Arg 4595	atg ctg gcg gag cga Met Leu Ala Glu Arg 4600	ccg gcg ggg Pro Ala Gly	39361
ctg gtg Leu Val 4605	ctg cgc aat gtg gcc Leu Arg Asn Val Ala 4610	aac gcc cgc acg gcg Asn Ala Arg Thr Ala 4615	gac gag gcg Asp Glu Ala	39406
gca cgg Ala Arg 4620	ctg gcg ctg ctg cgg Leu Ala Leu Leu Arg 4625	act ggc aac agc gtg Thr Gly Asn Ser Val 4630	ggc cgg ctg Gly Arg Leu	39451
cgg gca Arg Ala 4635	ctg ccg gcg gtg aca Leu Pro Ala Val Thr 4640	tcc tgg aat cca gag Ser Trp Asn Pro Glu 4645	cag ctc tgg Gln Leu Trp	39496
agg ctg Arg Leu 4650	gcg gag gca gcc gac Ala Glu Ala Ala Asp 4655	tac acg tgc cac gtc Tyr Thr Cys His Val 4660	acc tgg agc Thr Trp Ser	39541
gcc cag Ala Gln 4665	gac gag gaa ggc cgt Asp Glu Glu Gly Arg 4670	ttc gac gcg ctg ctc Phe Asp Ala Leu Leu 4675	atg gcc cgc Met Ala Arg	39586
gcg gct Ala Ala 4680	ggc tcc tcg cgc ccg Gly Ser Ser Arg Pro 4685	gcg gcc tgg ctg acg Ala Ala Trp Leu Thr 4690	cct ccc ccg Pro Pro Pro	39631
cct ccc Pro Pro 4695	cct cga ccc tgg aag Pro Arg Pro Trp Lys 4700	agc tac gcc aac cag Ser Tyr Ala Asn Gln 4705	ccc ctg gcg Pro Leu Ala	39676
gcc agc Ala Ser 4710	cgc cgg cgc acc ctc Arg Arg Arg Thr Leu 4715	gtg gga gtg ctg cga Val Gly Val Leu Arg 4720	agc cac ctc Ser His Leu	39721
gag cac Glu His 4725	aag ctg ccc gag tac Lys Leu Pro Glu Tyr 4730	atg gtg ccc tct tcc Met Val Pro Ser Ser 4735	ttc gtg ctg Phe Val Leu	39766
ctg gac Leu Asp 4740	gcg ctg ccg ctc aag Ala Leu Pro Leu Lys 4745	ccc acc ggc aag ttg Pro Thr Gly Lys Leu 4750	gag gtg gcg Glu Val Ala	39811
gca ctg Ala Leu 4755	cct ccc ccc gag ccc Pro Pro Pro Glu Pro 4760	gca caa gcc gag caa Ala Gln Ala Glu Gln 4765	acg cca ggc Thr Pro Gly	39856
cac ctc His Leu 4770	gcg ccg cgg acc ccc Ala Pro Arg Thr Pro 4775	acc gag cgc cgg ctt Thr Glu Arg Arg Leu 4780	gcc gag ctg Ala Glu Leu	39901

tgg cgg cgc gtg ctc ggc gtg cat cgc gtg ggg gtg gag gac aac 39946	
Trp Arg Arg Val Leu Gly Val His Arg Val Gly Val Glu Asp Asn	
4785 4790 4795	
ttc ttc cag ctg ggg ggc cac tcg ctt ctg gcc aca cgg ctc ttg 39991	
Phe Phe Gln Leu Gly Gly His Ser Leu Leu Ala Thr Arg Leu Leu	
4800 4805 4810	
tcg ctc atc cgc gga gag ctg ggg ctg gag ctg ccc ctg cgg gtg 40036	
Ser Leu Ile Arg Gly Glu Leu Gly Leu Glu Leu Pro Leu Arg Val	
4815 4820 4825	
ctc ttc gag cag ccc acc ctg gcg gcc atg gcg ggt tgt ctc gag 40081	
Leu Phe Glu Gln Pro Thr Leu Ala Ala Met Ala Gly Cys Leu Glu	
4830 4835 4840	
gct cag agc tgg agc acc cag gct ccc cac cct ccc tct gcc tcc 40126	
Ala Gln Ser Trp Ser Thr Gln Ala Pro His Pro Pro Ser Ala Ser	
4845 4850 4855	
atc gag gag gga gaa ctt tga gcgcgcggagc gttgctcgcc cacgccgcga 40177	
Ile Glu Glu Gly Glu Leu	
4860	
gcctgggtgt acgcctgtgg gtagagggcg agcgactgcg cttccaggct cccccagggg 40237	
tgatgacgcc cgagctgcag tcccgcctgg gaggcgctcg ccatgagctc atcgactgc 40297	
tacggcagct ccagccctcc tctcagggag gcagcctcct ggccccagtg gctcgaaacg 40357	
ggcggctggc gctctccttc gcccagcagc ggctgtgggt ccaggagcag ctccaccctg 40417	
aggccccggc caacaacctc acggggggccg tgggtgttcac cggcccgcgtg cacgtggccg 40477	
cgctgctcgg ggcgggtggct gcgctgggtga ggcgacacga agcgctgcgc accaccctgg 40537	
gcgaggaggg cggagtcccc tattcgctca tcggcgagcc gtggcagccg gcgctggagg 40597	
tggaagcgct accgggtgcc accgtgggcg agcggctgga gcaggcccga gaggtggccc 40657	
tggccgagtc ccggcgccgc ttgcacctgg agacggagcc ccacctgcgc gtgcgcctgc 40717	
tgcgcttggc cgagcagcag cacgtgctgg tgetcagcct ccaccacatc gcggcggatg 40777	
gagtcgggct gcaggtgctc gaacaggagc tggcggcgct ctacggcgcg ctgagcgctg 40837	
gggcccagacc gcggctacct cccctgcccc tgcaggtggc ggaccttgcc gactggcagc 40897	
gccgggtgggt ggagggagag gagtaccagg tccagctcgc ttactggcgg cgacagctcg 40957	
ccgggctcac gccctggag gtaccaggag atcatcctcg cccgcgtatc ccctccatga 41017	
gagggcgcgga ggtacgggcg cccctgctct ccgctccgca agcccagggtg ctgcgcgcgc 41077	
tggggcaagg cgagggcgcc accctctaca tgacgtcctt ggccgcccctg ggcgtgctgc 41137	
tgcaacggtg gacggggccag cacgatatgg cgggtgggcag cgccgcccgc aaccgcaacc 41197	
gccccgggctt ggagggcatc ctcggtcttc tcctcaacat cgtcctgctg cggctcgact 41257	
tgagaggccg cccccgcttc cgcgagttgc tgcgtcaggc acgcaggggtg tgtgtggagg 41317	
cctatgccca tcaggagttg cccttcgagc acctggtgga ggcgtccag ccaggcagcg 41377	

agcgcgggga cagctcgctc tatcgctgg cgctcgcggt gagcgacacg ccatggatgc 41437
ccggccacgg gctgaagctg gagggggtgc aagcgcagcc gctcgacttt ccccgggggg 41497
tgctggacct ggatctccac ctgtgggtgt atgacaccgg ggaggggctc accgggcggt 41557
tggagtacgc cgtggatctg tatgaggagc ccacggcaag gcggctcctg gagggcttcc 41617
gccaggtgct ggagggcggtg gtggaggcgc cggatcggcc cgtgccggag ctgccggtgc 41677
tgggagagca ggagcgccac caggtgctct ccgggtggaa ccgcacacag cggccctatc 41737
cgcgagaggc cagcgttcac ggcctcttcc agcagcgtgc cctccaggcc ccgcgggccg 41797
tggcagtggg gtatggcgag cgcagcctca cctacggcga gttggcggag cgggcccgcg 41857
ggctcgaca gggctctggg gcccgaggcg tgcggcgagg agacctgggc gcccttcggc 41917
tggagcgctc gccggagcag gtggagtcca tgctggccgt actccaagcc ggggccgcct 41977
acgtgccgct ggatccttcc taccocgtac agcgccagga gtcatgctc caggacagcg 42037
gcgcgcggtt gctcgtgcac agcgggccgc tgcccttcgc tcccagggc tgcgccacgc 42097
tggatctcca ggcattggcac cccgctccct ccgacggtgg cgagcccctg cccagtgca 42157
gcggagaaga tctggcgat gtcatctata cgtccggctc caccggccag cccaaggcg 42217
tggccgtgtg ccaccgggcc atgaccgcc tgggtgtgaa cacggactac gttcagctcg 42277
ggccagagga tcgggtcgcc caggcttcca atgcctcctt tgacgcggcc accttcgagg 42337
tctggggcgc gctgctcaat ggggcacggc tgggtgggct cgccacggag gaggccattc 42397
aggcgcgccg gctggccgag gtccctgcgc agcagcggt ctccgtgctc ttcgtcacca 42457
ccgccctctt caaccacgtg gcccgggagc aaccccaggc cttcagcacc ctgcgctacc 42517
tgctctttgg cggcgaggca gtggatgcct ccagtgtgag gcgggtgctg aaacagggcg 42577
ctccaggaca ccttctacac gtgtacggcc ccacggagaa caccaccttc tccaccgcct 42637
ggcggggtgga acacctggcc gagcaggcgc ataccgtgcc catgggtcat cccatcgcca 42697
actcgcggt gcacgtgctg gacgaagcgc tgcagccggt gccggtgggc gccatggggc 42757
aggtatacct gggaggtgat gggctggcgc tgggctactg gcgccacccc gaggccaccg 42817
cggagcgctt cgtgccggat ccgcacggcc tggagcccgg aggccggctc taccgcacgg 42877
ggatctggc gcggcggcag gcggatgggg ccgtgggtgtt cgctggccgc gtggacaggc 42937
aggtcaagct gcgaggcttc ccggtggagc cggcgagat cgagtcgcac ctgtgcgagc 42997
actcggaggt gtccgcggcg gtggtggagc tgcggggcga aggggccctg cggcggtgg 43057
tcgcttatgt cgtccccgc gccgggtggc gccaggtgc cgaggagctg cgcaccttc 43117
tcggagcag actgcccag tacatgctgc ccgccagctt ctccctgctg gaggcactgc 43177
ccctacccc caacggcaag gtggatcgct cggcgctacc cgagagcttc gaggaggcct 43237

cccgcgagca ggctccggtg gtacctccac gaggccccgt ggaggcgctg ctggtggaca 43297
tctggaggga ggtgctcggc acccaacggg tatccgtgca cgacgacttc ttcgacctgg 43357
gcggaacttc cctgctggcc acgcgggttg tctcccgtt gcgagaggcg ctgcaggtgg 43417
agttgccgct gcgcaccctc ttcgaggccc ctgagctctc ggcgctggcg gccaggtgg 43477
aggctcctgct cggccaccgg cagctccgcc cccctccgct cgtgcccgt gtacgtccac 43537
cagagctgcc cctctccttc gccagcagc ggctgtggtt cctccagcag ctgcgccac 43597
aaagcaccgc ctaccagatc ctggatgcgt ggcatgtgcg gggccgggtg gacgtggcg 43657
ccctggagcg cgccctggag cagctcgtgc gccggcatga ggcgctgcgc accactttcg 43717
agcctggcg cgacggcggtg cctcggcagc gcatccaagc gccggctccg gtcccgtgc 43777
ggcaggtgga cctgcgctcc catggcatcg ccgcgcggga ggaggccttg cgggtggatgc 43837
gcgagcaggc gctccggcca ctggagctgg acaaggggccc gctgctgcgc gtgtccctgc 43897
tgcgcttgga agaggcacag tccctgctct tccctgagct gcaccatata gtgggcatg 43957
gctggtcgct gagcggttg agccgagagc tgtcgacct gtacgaggcg gcgctccatg 44017
gcgctgagcc ggggttgcg ccgctgcccg tgcagtacgc ggacttcgcg ctgtggcagc 44077
gcggctggtt gcaagggccg gtgctgcggg aggagctgac ctggtggcg gagcggctcg 44137
cgcgctggc cccgctgcgc ctgcccgcgg accatgcccg cccggaagtg cagcgcttca 44197
acggcgcgac gtaccgcttc acgctcccgg gcgtacgggt gcaggcgctg cggcggtcg 44257
ggcatgagca cggggccaca ctcttcattg tgttgctgc cggttcaat gcgctgctgg 44317
cgcgctacac cgggcagacg gacatcgcca ttggcgcgcc catcgccaac cgcacacgtg 44377
gggaggtgga gggcctcatc ggcttcttcg tcaacaccct ggtgctgcgc accgcctgg 44437
aaggcaacct ctcttcctc gagctgttga ggcgggtgag ggagacaacg ctagaggcct 44497
acggccacca ggagctgcct ttcgagaggc tcgtcgagga gttgcagccc gagcgccagg 44557
ccaaccagaa cccgctggtg caggtgctgc tcgccctgca gaacgcgccc cgcgagccgt 44617
tgcggtggc gggcctggag gccgagcacc tggagtacct ggtggccacc acgcgcttcg 44677
acctcgagct gcacctgtgg gaggaggaag aggggctgag ctgcatcgcc gtgtatgacc 44737
gcgcgctgta cggcgcgggg acggtggagc ggctggtggg cgctggtgc acgctgctgg 44797
agggcggtggc cgagctcccg gcgcgccgcg tggcgagct gccctggtg cctgcccggg 44857
agctgcgaca ggttccgccc cctcccccg ccgcggcgat cgaaacaagc atcggcgccc 44917
gcttctcgga ggtggcgcg cggcagccag gcgccacggc cgtcaccaca ggcgggcggc 44977
acctaccta tgcagagctg gaggagcgct cggaacgcct gggcgctac ctggcctggc 45037
tgggcgctgc cgcgggcgac agagtgggt tggccaccga gcgcagctg gagcggatca 45097
tcagcctgct aggcattctc aaggcgggag ccgcctacgt accgctggat gtgcgccagc 45157

cagcgcgctcg gctgagcctg ctggtgcagg ccgcggggcgt gcgcaccgtc atcgccgagg 45217
agcaggctcg cacggtgctg tcaggcctgg ggcagccact gaccctggtc gatgccgccc 45277
aggagccggc ctccgcgcag caggtccccg cgctcgggccc ggagcgctcg ctgggcgggg 45337
acatgctcgc ctacgtgctt ttcacctcgg gctccaccgg cgagcccaag ggcgtctgca 45397
ttccccaccg tgccgtgctg cggctgatcc atgagccctc ctacgtccag ctctcgccgc 45457
gggaggtcat gctgcactac gcaccgctgg agttcgatgc ctccaccttc gaggtctggg 45517
gggcgctgct caacggggca aggctcgtgc tggctccctcc ggaacagcag tcgctcgaga 45577
gcctgggaca ggagctgagc acccagggcg tcaccgtgct gtggctcacc gccgggctct 45637
tccggctcat ggtggaggag cagctcaaga gcctgcgcgg cgtgcgccag ctgctcgccg 45697
gaggggacgt gctgcccatt ccgcaggtac gccggctgcg cgaggccctc cccgagtgc 45757
agctcatcaa cggctacggc cccaccgaga gctgcacctt cacctgctgt caccgcgtgg 45817
gcagcccccga ggagctgggc ggctccgtgc ccacggcac ccccatcgac ctgggctggg 45877
tgagcggtgt ggacgagcgg ctccagcccg tgccggacgg agccccgggc gagctgctcg 45937
tgggcgggccc cgggctggct tgggggtacc tgcaacaccc ggagctcacc gccgagcgct 45997
tcacccccga tccgctcagc cggacgccgg gggctcgcgt ctatcggacg ggagacctgg 46057
tgccggcgag ggaggatggg acgctggagt tcctggggccg cgtggaccac cagctcaagg 46117
tacgcggctt ccgcacgag ccgggcgagg tggaggccgc ggtgctcacc caccggcg 46177
tgacgtccgc ggtgggtggtg gggcgcgagg gccccggcg caaggagctg gtctgctacg 46237
ccgtgcccg ggtggagagc tcggagcagg gctcgcagca ggagcagcgg ctggtacacg 46297
agtgggagtc cgtgttcgac gggcacatgt accgcgaagc ccccgctgggc ggagagccga 46357
ccttcaacat cgtcggctgg aagagcagct acaccggcca acccgctggc gtggaggaga 46417
tgagggactg gctgcgccac cgggtggagc gcgtgcgcgg gctcaggcca cgacggatcc 46477
tcgaagtggg ctgtggtacg gggctgatgc tctttgccct gctgccgcac tgcgagcgct 46537
acgtgggcac ggatttctcc ccggcggcgt tggactacgt gcgtcgctac ctccccccgg 46597
agcaccggg ccgcgtggag ctgctacacc gcacggcgga cgagtggagc ggcgtggcg 46657
ccggctcctt cgacgcggtg ctgctcaatt cgggtggtgca gtacttcccc tcgcaggagt 46717
acctgcggca ggtgctggcc cgctgcgtgg aggcctgga ggacgggggc ttcgtcttcg 46777
tgggggatgt gcggagcttg ccgctgctgg agtccttcca cgcctcgggtg gagctggagc 46837
ggggcgcccc ctccatgccg ctggaggcgt ggcgcgagcg ggtgcggcgt gcggtgctgg 46897
aggacaacga gctggtggtg gaccggcac tcttcgtggc gctggcccat cagcatccac 46957
gggtgagcca cgtggatatc gagctgacgc gcggcacgca cccgaatgag atggcacgct 47017

tccgctacaa cgccgtgctc cacatcgggc cgcgacgcc gccaccggcc tccgaggtgc	47077
cgtgggtgga ctggtccacg caaggactga gcctcgatgc actgcgggct cggctccggc	47137
agggcccgcc agggccactg ggggtggcag gcattcccaa tgcgcgagtg ctgcccgcgg	47197
tccggggcgc cgaggcgctg ggcagcacag gcagcgcgcg gcgggtggag gagctgcgcc	47257
ggcgcctgtc ccaaccggc acagggggcc aggatccaga ccccttcttg caactggccg	47317
agtcgctcgg ctacacggcg gcagtgcgct ggtctccttg gcggaggac ggggccttcg	47377
acgtgctctt cctcccgcg acgccaggga tgcaccgcg gtggctcggc cccacgccgc	47437
tcaaccgcac gccgccccca gcctccgcgc gtctgtcctc ggagcctcgg cgggccagcc	47497
tctcactgag gctgggctcc gcgctgcggg cccacctgca gacctatcta ccagacttca	47557
tggtgccctc acgcttcgtg gtgctgcagt cctgcgcgt cactcccaat ggcaaggtgg	47617
accgcgcggc cttgcccgtc cccgactcgc gtcgggtgga atctgccccg ctggtgcccc	47677
ccagcaacga gctggagcgg gtgctggcgc aggtgtggaa ggaagtgtg ggactggagg	47737
aagtgcgcgg ggaggacaac ttcttcgacg tgggagggca ctgctgctg ctggcgagg	47797
tgtgcagccg gctggaggcc cgactgggccc ggcggtcga actggtgacc ctcttcgcgt	47857
actcctccat tgctgccctg gccgagcacc tgcaggcccc ccaggagctg gcagccgccg	47917
aggcgcaggt gcagcgcagt gccacggagc gcgccttgct ccagcaacag gcagcccaac	47977
gccgcccgggc cggttcgaag cggggagctc cca atg	48028
Met Thr Pro Ser Gly Asp	4865 4870
gaa gcg ctg cag tct tcc atc gcc ctg gta ggc atg gcc ggc cgc	48073
Glu Ala Leu Gln Ser Ser Ile Ala Leu Val Gly Met Ala Gly Arg	4875 4880 4885
ttc ccg ggg gcc ccg gac gtg gag tcc ttc tgg cgc aat ctg gtg	48118
Phe Pro Gly Ala Pro Asp Val Glu Ser Phe Trp Arg Asn Leu Val	4890 4895 4900
gcc gga gtg gag tcc att tcc ttc ttc tcc gag gaa gag ctg aga	48163
Ala Gly Val Glu Ser Ile Ser Phe Phe Ser Glu Glu Glu Leu Arg	4905 4910 4915
cag gcg ggc gtg tcc gag cag ata cgt cgg cgc ccc gag tac gtg	48208
Gln Ala Gly Val Ser Glu Gln Ile Arg Arg Arg Pro Glu Tyr Val	4920 4925 4930
ccc gcc aag ggt gtg ctg gaa gac ctg gag ctg ttc gac gcg ggc	48253
Pro Ala Lys Gly Val Leu Glu Asp Leu Glu Leu Phe Asp Ala Gly	4935 4940 4945
ttc ttc ggc tac tcg ccg cgc gag gcg agc cac ctg gat ccg cag	48298
Phe Phe Gly Tyr Ser Pro Arg Glu Ala Ser His Leu Asp Pro Gln	4950 4955 4960
caa cgc ctg ctg ctg gag tgc tcg tgg gag gcg ctg gag gat gcg	48343
Gln Arg Leu Leu Leu Glu Cys Ser Trp Glu Ala Leu Glu Asp Ala	4965 4970 4975

ggc ctt cgc ccg gat Gly Leu Arg Pro Asp 4980	cag ctc cca ggc tgg Gln Leu Pro Gly Trp 4985	gtg ggc gtc tac gtg Val Gly Val Tyr Val 4990	48388
ggc gcc ggg gac acc Gly Ala Gly Asp Thr 4995	agc tac cgc ttc cag Ser Tyr Arg Phe Gln 5000	ctg ctg cgc gga cac Leu Leu Arg Gly His 5005	48433
ggg gac ccg ctc agc Gly Asp Pro Leu Ser 5010	ggc tcc aag gat Gly Ser Lys Asp 5015	gtg gcg ggc ttc ttt ggc Val Ala Gly Phe Phe Gly 5020	48478
aac tac ccg gac ttt Asn Tyr Pro Asp Phe 5025	ctc gcc acg cgc gtg Leu Ala Thr Arg Val 5030	gcc tac aag ctc aat Ala Tyr Lys Leu Asn 5035	48523
ctg cgc ggc ccc gcg Leu Arg Gly Pro Ala 5040	ctc ggc atc cac acc Leu Gly Ile His Thr 5045	gcg tgc tcc acc tcg Ala Cys Ser Thr Ser 5050	48568
ctc gtc tcc atc aac Leu Val Ser Ile Asn 5055	atg gcg tgc agc gcg Met Ala Cys Ser Ala 5060	ctg cgc ggc ttc gag Leu Arg Gly Phe Glu 5065	48613
tgt gac atg gcg ctc Cys Asp Met Ala Leu 5070	gcg ggc ggc gtc tcg Ala Gly Gly Val Ser 5075	ctg cgg ctg ccc gcg Leu Arg Leu Pro Ala 5080	48658
cgc tcc ggc tac ctc Arg Ser Gly Tyr Leu 5085	tac gag gag ggg gga Tyr Glu Glu Gly Gly 5090	gtc gcc tcg aag gac Val Ala Ser Lys Asp 5095	48703
ggg cac tgc cgc ccc Gly His Cys Arg Pro 5100	ttc gac gcg cgt gct Phe Asp Ala Arg Ala 5105	acc ggc acc gtc acg Thr Gly Thr Val Thr 5110	48748
ggg gat ggc gtg ggc Gly Asp Gly Val Gly 5115	gtg gtg gtg ctc aag Val Val Val Leu Lys 5120	cgc ctg gag gac gcg Arg Leu Glu Asp Ala 5125	48793
ctg aag gcc agg gat Leu Lys Ala Arg Asp 5130	ccc atc cac gcg gtg Pro Ile His Ala Val 5135	atc cgc ggc tgg gcg Ile Arg Gly Trp Ala 5140	48838
ctc aac aac gac ggc Leu Asn Asn Asp Gly 5145	gcc tcc cgg gct ggc Ala Ser Arg Ala Gly 5150	ttc acc gcg ccc agc Phe Thr Ala Pro Ser 5155	48883
gtg gag ggc caa tcc Val Glu Gly Gln Ser 5160	gag gtg atc gcc ctg Glu Val Ile Ala Leu 5165	gcg cac gcc gcg gcg Ala His Ala Ala Ala 5170	48928
ggc atc agc gca cgc Gly Ile Ser Ala Arg 5175	gac atc acc tac gtg Asp Ile Thr Tyr Val 5180	gag gct cac ggc acc Glu Ala His Gly Thr 5185	48973
ggc act ccc ctg gga Gly Thr Pro Leu Gly 5190	gac ccg atc gaa gtg Asp Pro Ile Glu Val 5195	gcg gcg ctc acc cgc Ala Ala Leu Thr Arg 5200	49018
gcc ttc cgg gcg cat Ala Phe Arg Ala His 5205	acc gct gac acc gcc Thr Ala Asp Thr Ala 5210	ttc tgt acc ctc ggt Phe Cys Thr Leu Gly 5215	49063

5205	5210	5215	
gcg gtc aaa tcc aac atc gga cac ctg gat Ala Val Lys Ser Asn Ile Gly His Leu Asp 5220	gcg gcg gca gga gtg Ala Ala Ala Gly Val 5225	49108	5230
gcc ggg gtg atc aag acg gtg cag gcc ctg Ala Gly Val Ile Lys Thr Val Gln Ala Leu 5235	cgc cac cgg ctc atc Arg His Arg Leu Ile 5240	49153	5245
cct ccc acg ctc cac ttc gag cgg ccc aat Pro Pro Thr Leu His Phe Glu Arg Pro Asn 5250	ccc gcg ctc cac ctg Pro Ala Leu His Leu 5255	49198	5260
gag cag agc ccc ttc ttc gtc aac acc cag Glu Gln Ser Pro Phe Phe Val Asn Thr Gln 5265	ccg ctc ccg tgg gag Pro Leu Pro Trp Glu 5270	49243	5275
agc ccg cgc ggg cct cgg ctg gcg ggg gtg Ser Pro Arg Gly Pro Arg Leu Ala Gly Val 5280	agc tcc ttt ggc att Ser Ser Phe Gly Ile 5285	49288	5290
ggc ggc acc aac gcc cac acg ctc ttc gaa Gly Gly Thr Asn Ala His Thr Leu Phe Glu 5295	gag gct ccg ccg ccg Glu Ala Pro Pro Pro 5300	49333	5305
ccc gcc agc ggc ccc acc cgt ccc aac cag Pro Ala Ser Gly Pro Thr Arg Pro Asn Gln 5310	gtg ctg ctg ctc tcc Val Leu Leu Leu Ser 5315	49378	5320
gcc cga agc acc agc gcg ctc gag cac atc Ala Arg Ser Thr Ser Ala Leu Glu His Ile 5325	gca gga cgg ctc gcg Ala Gly Arg Leu Ala 5330	49423	5335
gca cac ctg cgc cgc cac ccg gac ctg gaa Ala His Leu Arg Arg His Pro Asp Leu Glu 5340	ctg gcg gac gtg gcc Leu Ala Asp Val Ala 5345	49468	5350
ttt acc ctc cag gtg ggg cga gcc cgc ttc Phe Thr Leu Gln Val Gly Arg Ala Arg Phe 5355	ccc tac cgg cgc gcc Pro Tyr Arg Arg Ala 5360	49513	5365
ctc acc tgt cgc acc ctg gcc gag gcc atg Leu Thr Cys Arg Thr Leu Ala Glu Ala Met 5370	gag cgc ctg gag gct Glu Arg Leu Glu Ala 5375	49558	5380
ccc gag ccg cgg cca ccc gag ccg ctt gct Pro Glu Pro Arg Pro Pro Glu Pro Leu Ala 5385	cac gag gga gag ccg His Glu Gly Glu Arg 5390	49603	5395
ccc cct ctc gtc atg ctc ttt cca ggg cag Pro Pro Leu Val Met Leu Phe Pro Gly Gln 5400	ggc act ccg ctg gtc Gly Thr Pro Leu Val 5405	49648	5410
ggc acg gcc cgt gcg ctc cat gaa tcc gag Gly Thr Ala Arg Ala Leu His Glu Ser Glu 5415	ccc acc ttc cgc cag Pro Thr Phe Arg Gln 5420	49693	5425
gcg gtg gag cag tgt gcc cgg ctc ctg agg Ala Val Glu Gln Cys Ala Arg Leu Leu Arg 5430	cag acg ctg ggg ctg Gln Thr Leu Gly Leu 5435	49738	5440
gat gtc agg gag gtg ctc ttc ccc tcc gcg	gag cag gag gag cag	49783	

Asp Val Arg Glu Val	Leu Phe Pro Ser Ala	Glu Gln Glu Glu Gln	
5445	5450	5455	
gca cgc cgg ctg gcg	gcg cag acg cgg gtg	gcc cag ccc gcg ctc	49828
Ala Arg Arg Leu Ala	Ala Gln Thr Arg Val	Ala Gln Pro Ala Leu	
5460	5465	5470	
ttc acc ctg gag tat	gcc ctg gcg cag acc	tgg ctg ggc tgg ggg	49873
Phe Thr Leu Glu Tyr	Ala Leu Ala Gln Thr	Trp Leu Gly Trp Gly	
5475	5480	5485	
ctc caa ccg cag gcg	ctg gca ggc cat agc	ctg ggc gag ctg gtg	49918
Leu Gln Pro Gln Ala	Leu Ala Gly His Ser	Leu Gly Glu Leu Val	
5490	5495	5500	
gcg gcg tgc ctc gcg	ggg gtc ttc tcc ctg	gaa gat gca ctt cag	49963
Ala Ala Cys Leu Ala	Gly Val Phe Ser Leu	Glu Asp Ala Leu Gln	
5505	5510	5515	
ctc gtg gca gcc cgg	gga cag ctc atg cag	gga tgc cca ccg ggg	50008
Leu Val Ala Ala Arg	Gly Gln Leu Met Gln	Gly Cys Pro Pro Gly	
5520	5525	5530	
gcc atg ctg gcc gtg	ccg ctg ccg gag gcc	gag ctg gcg gcg ctg	50053
Ala Met Leu Ala Val	Pro Leu Pro Glu Ala	Glu Leu Ala Ala Leu	
5535	5540	5545	
ctg ggc agc gag ctg	tgt atc gcc gcg gtg	aac ggg cca ccg gcc	50098
Leu Gly Ser Glu Leu	Cys Ile Ala Ala Val	Asn Gly Pro Arg Ala	
5550	5555	5560	
tgc gtg gcc tcg ggt	cct cta ccg gcg gta	gag gcg ctc acc gcc	50143
Cys Val Ala Ser Gly	Pro Leu Pro Ala Val	Glu Ala Leu Thr Ala	
5565	5570	5575	
gcg ctg gag agc cgg	ggc gtg tcc agc cgg	cgc ctg gag act tcc	50188
Ala Leu Glu Ser Arg	Gly Val Ser Ser Arg	Arg Leu Glu Thr Ser	
5580	5585	5590	
cat gcc ttc cac tct	gcc tcc atg gag gca	tgc cag ggg cca ctc	50233
His Ala Phe His Ser	Ala Ser Met Glu Ala	Cys Gln Gly Pro Leu	
5595	5600	5605	
acc acg ctg ctg cgg	cgc atg cgc ctg cag	gcc cct cgc ctc ccc	50278
Thr Thr Leu Leu Arg	Arg Met Arg Leu Gln	Ala Pro Arg Leu Pro	
5610	5615	5620	
tgt gtc tcc ggc ctc	acc ggc cgg tgg ctc	acc ggc gag gag gcc	50323
Cys Val Ser Gly Leu	Thr Gly Arg Trp Leu	Thr Gly Glu Glu Ala	
5625	5630	5635	
acg gag cct acg tac	tgg gcc cgc cag ctc	cgc gag ccg gta cgt	50368
Leu Glu Pro Thr Tyr	Trp Ala Arg Gln Leu	Arg Glu Pro Val Arg	
5640	5645	5650	
ttc tcc gag gcc ctc	gag acg ctc tgg agc	ctc aag gag cca gtg	50413
Phe Ser Glu Ala Leu	Glu Thr Leu Trp Ser	Leu Lys Glu Pro Val	
5655	5660	5665	
acc gag gag gta ggc	ccc ggc acc acg ctc	acc gcg ctt gcg cgg	50458
Leu Leu Glu Val Gly	Pro Gly Thr Thr Leu	Thr Ala Leu Ala Arg	
5670	5675	5680	

cgg cac ccc acc cgc	ccc gcc cgc acc cag	gag gtg gcg agc ctt	50503
Arg His Pro Thr Arg	Pro Ala Arg Thr Gln	Glu Val Ala Ser Leu	
5685	5690	5695	
ccc gtg caa ccg gac	acg gcc gtg ccg tgc	att gaa gag gcc gtg	50548
Pro Val Gln Pro Asp	Thr Ala Val Pro Cys	Ile Glu Glu Ala Val	
5700	5705	5710	
ggt gag ctg tgg cag	gcc ggc ctg gag ctg	gac tgg agc gcc ctg	50593
Gly Glu Leu Trp Gln	Ala Gly Leu Glu Leu	Asp Trp Ser Ala Leu	
5715	5720	5725	
cac gcc gcg cca cgc	cac cgc gca cac ctt	cca ccc tac ccc ttc	50638
His Ala Ala Pro Arg	His Arg Ala His Leu	Pro Pro Tyr Pro Phe	
5730	5735	5740	
gag cgc cag cgc tac	tgg att gaa ccg gag	gct gcc ccg cag ccg	50683
Glu Arg Gln Arg Tyr	Trp Ile Glu Pro Glu	Ala Ala Pro Gln Pro	
5745	5750	5755	
cgc gcg cag cag ccg	acg ccc gcg tcc ctt	gtg ccc ccg gag cag	50728
Arg Ala Gln Gln Pro	Thr Pro Ala Ser Leu	Val Pro Pro Glu Gln	
5760	5765	5770	
ccc tcg cgc gag gcc	ctg gaa gac tgg ttc	tac gtg ccc acc tgg	50773
Pro Ser Arg Glu Ala	Leu Glu Asp Trp Phe	Tyr Val Pro Thr Trp	
5775	5780	5785	
gag cag gca cct gcc	acg agc ggg ggg gga	cag ccc ctg gcg ggc	50818
Glu Gln Ala Pro Ala	Thr Ser Gly Gly Gly	Gln Pro Leu Ala Gly	
5790	5795	5800	
ccg gtg ctc gcc ttc	atg gac tcc tcg ggc	ctg gcc gag cag gtg	50863
Pro Val Leu Ala Phe	Met Asp Ser Ser Gly	Leu Ala Glu Gln Val	
5805	5810	5815	
ctg gcc gcg ctg tgg	ccc gcg gac tcc ggc	gcg ctc ctc acc cgc	50908
Leu Ala Ala Leu Trp	Pro Ala Asp Ser Gly	Ala Leu Leu Thr Arg	
5820	5825	5830	
gtc gag ccg gcc ggg	cac tat gag caa ctc	agt gag cac gcc ttc	50953
Val Glu Pro Ala Gly	His Tyr Glu Gln Leu	Ser Glu His Ala Phe	
5835	5840	5845	
cgc ctc cgc ccc gag	agc gag gag gac tgg	gac gcg ctc ttc cag	50998
Arg Leu Arg Pro Glu	Ser Glu Glu Asp Trp	Asp Ala Leu Phe Gln	
5850	5855	5860	
gcg ctc cag tcc cag	ggc cgg ctc ccc cgt	cgc atc ctc cat gcc	51043
Ala Leu Gln Ser Gln	Gly Arg Leu Pro Arg	Arg Ile Leu His Ala	
5865	5870	5875	
tgg gct ctc acc gcg	gag ccc ggc ccc tgt	acc ccg gac ggg gaa	51088
Trp Ala Leu Thr Ala	Glu Pro Gly Pro Cys	Thr Pro Asp Gly Glu	
5880	5885	5890	
gcg gtg ctg gaa cag	ggc ttc ttc agc ctg	ctg cgg ctg gcc cgt	51133
Ala Val Leu Glu Gln	Gly Phe Phe Ser Leu	Leu Arg Leu Ala Arg	
5895	5900	5905	
gcg ctc ggc cgc cac	gcg cct gag cgc cct	gta cag ctc gag gtg	51178
Ala Leu Gly Arg His	Ala Pro Glu Arg Pro	Val Gln Leu Glu Val	
5910	5915	5920	

ctg tcc agc ttc gtc	cac gcg gtg ggc ccg	cgg gag ccg ctc gag	51223
Leu Ser Ser Phe Val	His Ala Val Gly Pro	Arg Glu Pro Leu Glu	
5925	5930	5935	
ccg ctc aag gcc acg	ctg ctg ggg gcc tgt	gcg gta ctt ccc ctg	51268
Pro Leu Lys Ala Thr	Leu Leu Gly Ala Cys	Ala Val Leu Pro Leu	
5940	5945	5950	
gag tac cct cac gtg	cag tgc cgc acc att	gac gtg ccg ccc ggg	51313
Glu Tyr Pro His Val	Gln Cys Arg Thr Ile	Asp Val Arg Pro Gly	
5955	5960	5965	
agc gag cca ccg gaa	gtg ctg gtg agg agc	ctg gcc gcc gag ctg	51358
Ser Glu Pro Arg Glu	Val Leu Val Arg Ser	Leu Ala Ala Glu Leu	
5970	5975	5980	
gcc gct ccc atg gga	gag agc ccg gtg gcc	tgg cgc gac ggg cag	51403
Ala Ala Pro Met Gly	Glu Ser Pro Val Ala	Trp Arg Asp Gly Gln	
5985	5990	5995	
cgg tac gtg cgc cgc	gcc acc ccg cag agg	ctc gag gcc tcg ccg	51448
Arg Tyr Val Arg Arg	Ala Thr Arg Gln Arg	Leu Glu Ala Ser Arg	
6000	6005	6010	
ccc ctg cgc agc ctt	cgg gag cga gcc gtg	tac ctg gta gca gga	51493
Pro Leu Arg Ser Leu	Arg Glu Arg Gly Val	Tyr Leu Val Ala Gly	
6015	6020	6025	
ggg ctg ggt gga atc	ggc ctg gtg ctg gcg	cga gcg ctg gcc cag	51538
Gly Leu Gly Gly Ile	Gly Leu Val Leu Ala	Arg Ala Leu Ala Gln	
6030	6035	6040	
cgg gcg cgt gct ccg	ctg gcg cta ctc acc	cac tcg ccc ttc cct	51583
Arg Ala Arg Ala Arg	Leu Ala Leu Leu Thr	His Ser Pro Phe Pro	
6045	6050	6055	
ccc cga gag cag tgg	gag cag tgg ctg gag	gaa gcc ccg gcg cac	51628
Pro Arg Glu Gln Trp	Glu Gln Trp Leu Glu	Glu Ala Pro Ala His	
6060	6065	6070	
ccg gag ccg gcg tgg	cgg agc gag gcc gac	ccc tcg gag cgc cgc	51673
Pro Glu Pro Ala Trp	Arg Ser Glu Ala Asp	Pro Ser Glu Arg Arg	
6075	6080	6085	
agg acc cag cac cgc	atc cgc tgc ctg ctg	gag ctg gag cag ctc	51718
Arg Thr Gln His Arg	Ile Arg Cys Leu Leu	Glu Leu Glu Gln Leu	
6090	6095	6100	
ggc gcc gag gtg cag	gtg tac acg gcg gac	gtg gcg gag gag gcc	51763
Gly Ala Glu Val Gln	Val Tyr Thr Ala Asp	Val Ala Glu Glu Ala	
6105	6110	6115	
gcc gtg cgc tcg gtg	gtg gag cag gtg cac	gcc cgc tgg ggg aag	51808
Ala Val Arg Ser Val	Val Glu Gln Val His	Ala Arg Trp Gly Lys	
6120	6125	6130	
atg gag gcc gtg ctg	cac gcc gcc gcc acg	ttc gac gac gcc gtc	51853
Ile His Gly Val Leu	His Ala Ala Ala Thr	Phe Asp Asp Gly Val	
6135	6140	6145	
acc cag ctc cgc acg	cac gag cag tcc tcg	cgg gcc ctg cgc acc	51898
Ile Gln Leu Arg Thr	His Glu Gln Ser Ser	Arg Ala Leu Arg Thr	

6150	6155	6160	
aag gtg cgg ggc agc Lys Val Arg Gly Ser 6165	atg gtg ctg cac gag Met Val Leu His Glu 6170	gtg ctg gcg agc gag Val Leu Ala Ser Glu 6175	51943
ggg ttg gat tgg ttc Gly Leu Asp Trp Phe 6180	gcc ctg tgc tcc tcg Ala Leu Cys Ser Ser 6185	ctg gcg tcg gcg ctg Leu Ala Ser Ala Leu 6190	51988
ggc tca ttc ggc cag Gly Ser Phe Gly Gln 6195	gcg gac tac tgc gcg Ala Asp Tyr Cys Ala 6200	gcc aat gcc ttc cag Ala Asn Ala Phe Gln 6205	52033
gat gcg tat gcg cac Asp Ala Tyr Ala His 6210	cac ctg cgc cgg cag His Leu Arg Arg Gln 6215	ggc ttc acg gga gcg Gly Phe Thr Gly Ala 6220	52078
ctg gcg ctg gac tgg Leu Ala Leu Asp Trp 6225	ggc acg tgg aga gat Gly Thr Trp Arg Asp 6230	acg ggg gcg gcc atg Thr Gly Ala Ala Met 6235	52123
cgg ctg gtg gca cgc Arg Leu Val Ala Arg 6240	acc cga cgg ggg ggc Thr Arg Arg Gly Gly 6245	cat gag aag ccg ccc His Glu Lys Pro Pro 6250	52168
acg ccg ctc acc cac Thr Pro Leu Thr His 6255	cca ctc ttc gat tgc Pro Leu Phe Asp Cys 6260	gaa cag cgc gag ccg Glu Gln Arg Glu Pro 6265	52213
ggc ggg acg cac tgg Gly Gly Thr His Trp 6270	ctc ggc ctc acg ctg Leu Gly Leu Thr Leu 6275	agg ggt gga gag gac Arg Gly Gly Glu Asp 6280	52258
tgg gtg gtg gac gag Trp Val Val Asp Glu 6285	cac cgg ctg caa ggg His Arg Leu Gln Gly 6290	gtg cca aca ctg ccc Val Pro Thr Leu Pro 6295	52303
ggg gtg gcc tac ctg Gly Val Ala Tyr Leu 6300	gag ctg gcg cgg gcg Glu Leu Ala Arg Ala 6305	gcg tgt gcc cag gcg Ala Cys Ala Gln Ala 6310	52348
ctg ggg gcc gag gcg Leu Gly Ala Glu Ala 6315	gtg gag ctg gcg gag Val Glu Leu Ala Glu 6320	ctg ttg ctg ctg gag Leu Leu Leu Leu Glu 6325	52393
cct ctg acg gta cct Pro Leu Thr Val Pro 6330	cga ggc gaa tcc agg Arg Gly Glu Ser Arg 6335	cag gtc cgg gtg gtg Gln Val Arg Val Val 6340	52438
ctc cag ccc gag ggg Leu Gln Pro Glu Gly 6345	cag gct cat gcc ctg Gln Ala His Ala Leu 6350	cgg gtg gag agc cgg Arg Val Glu Ser Arg 6355	52483
tcg gag gaa gcg cgg Ser Glu Glu Ala Arg 6360	gga tgg aat gag cat Gly Trp Asn Glu His 6365	gcg cgg ggc cgg gtg Ala Arg Gly Arg Val 6370	52528
cgc gcc gtg cct cgg Arg Ala Val Pro Arg 6375	ctg gcc gag cgc atc Leu Ala Glu Arg Ile 6380	cag ccc gag ctg ctg Gln Pro Glu Leu Leu 6385	52573
cgc gcc gcc tgc gaa	cac gag cag ccc gtg	ccc gga gag ccc cag	52618

Arg Ala Ala Cys Glu	His Glu Gln Pro Val	Pro Gly Glu Pro Gln	
6390	6395	6400	
gag caa ggc cct gtc	cac gcc gga gca cgc	tgg cat ggc ctc ttc	52663
Glu Gln Gly Pro Val	His Ala Gly Ala Arg	Trp His Gly Leu Phe	
6405	6410	6415	
cag tgg gtt cgc cgc	ggc cct cgc cag gcc	ctc gcc cag ctc gcg	52708
Gln Trp Val Arg Arg	Gly Pro Arg Gln Ala	Leu Ala Gln Leu Ala	
6420	6425	6430	
cta ccg gag ccc ttc	cac ggt gac ctc gag	cgc ttc gag ctc cac	52753
Leu Pro Glu Pro Phe	His Gly Asp Leu Glu	Arg Phe Glu Leu His	
6435	6440	6445	
ccg gcg ctc atg gac	atg gcc acc agc ttc	gcc att ccc ggc gga	52798
Pro Ala Leu Met Asp	Met Ala Thr Ser Phe	Ala Ile Pro Gly Gly	
6450	6455	6460	
gtg ccc tgg ctc gcc	ttt ggc tat gag cgc	gtg ctc att cac ggt	52843
Val Pro Trp Leu Ala	Phe Gly Tyr Glu Arg	Val Leu Ile His Gly	
6465	6470	6475	
ccg ctg cca ccg cag	gtg ctc agc cac gtg	agc ctg ccc gag gag	52888
Pro Leu Pro Pro Gln	Val Leu Ser His Val	Ser Leu Pro Glu Glu	
6480	6485	6490	
tcg cag gct ggc gcg	caa caa ctc cgg cta	cag gtg cgg ctg ctg	52933
Ser Gln Ala Gly Ala	Gln Gln Leu Arg Leu	Gln Val Arg Leu Leu	
6495	6500	6505	
gat ctg gag ggc tgg	gag cgg gtg cgg atc	gat ggg tac ctg ctg	52978
Asp Leu Glu Gly Trp	Glu Arg Val Arg Ile	Asp Gly Tyr Leu Leu	
6510	6515	6520	
cgg ccg ctg aag ccc	agc gac gcc agt gtc	gag ccg gcg gcg ccc	53023
Arg Pro Leu Lys Pro	Ser Asp Ala Ser Val	Glu Pro Ala Ala Pro	
6525	6530	6535	
aac gtg gaa gtc gcc	gtg ggg acg ccg ggg	ctg ctg gag agc ctt	53068
Asn Val Glu Val Ala	Val Gly Thr Pro Gly	Leu Leu Glu Ser Leu	
6540	6545	6550	
ggc ctg cgc cgc tgc	acc cgc ccc gcc ccc	ggg ccg cgc cag gtg	53113
Gly Leu Arg Arg Cys	Thr Arg Pro Ala Pro	Gly Pro Arg Gln Val	
6555	6560	6565	
gag atc gag gtg gag	gcc gct ggg ctc aac	ttc ctg gac gtg ctg	53158
Glu Ile Glu Val Glu	Ala Ala Gly Leu Asn	Phe Leu Asp Val Leu	
6570	6575	6580	
ggc gcg ctg ggg atg	atg ccg gca ttg gag	gcg gag gag agc gta	53203
Gly Ala Leu Gly Met	Met Pro Ala Leu Glu	Ala Glu Glu Ser Val	
6585	6590	6595	
ctg ggg cgc gag tgc	tcc gga cgc att gcc	gcc gtg ggc gag ggc	53248
Leu Gly Arg Glu Cys	Ser Gly Arg Ile Ala	Ala Val Gly Glu Gly	
6600	6605	6610	
gtc agc ggg ctg cgc	gtg ggg gac gag gtg	ctg gcg gtg gcc cca	53293
Val Ser Gly Leu Arg	Val Gly Asp Glu Val	Leu Ala Val Ala Pro	
6615	6620	6625	

ggc tgc ttc cgc tcc	tac gtg ctg gtg gat	gag agc cag gtg gtg	53338
Gly Cys Phe Arg Ser	Tyr Val Leu Val Asp	Glu Ser Gln Val Val	
6630	6635	6640	
cgc agg ccc gcc tcg	ctg ggg ctc gcc gag	ggg gcg gcc cag atg	53383
Arg Arg Pro Ala Ser	Leu Gly Leu Ala Glu	Gly Ala Ala Gln Met	
6645	6650	6655	
gtg ccg ttc gcc acg	gcg tac ttc gcc ctg	cac acc gtg ggc cgg	53428
Val Pro Phe Ala Thr	Ala Tyr Phe Ala Leu	His Thr Val Gly Arg	
6660	6665	6670	
ttg cgg cgc ggc gag	cgc atc ctc atc cac	gcc gcg gcc gga ggg	53473
Leu Arg Arg Gly Glu	Arg Ile Leu Ile His	Ala Ala Ala Gly Gly	
6675	6680	6685	
ctg ggt ctg gcc gcc	gtt cag ctc gca tcc	cgg acc ggg gcg gag	53518
Leu Gly Leu Ala Ala	Val Gln Leu Ala Ser	Arg Thr Gly Ala Glu	
6690	6695	6700	
ata ttg gcc acc gcg	ggc agt gag cag aag	cgc gag tac ctg cgc	53563
Ile Leu Ala Thr Ala	Gly Ser Glu Gln Lys	Arg Glu Tyr Leu Arg	
6705	6710	6715	
tcg ctc ggc att gct	cac gtg ctg gac tcg	cgc agc acc tcc ttc	53608
Ser Leu Gly Ile Ala	His Val Leu Asp Ser	Arg Ser Thr Ser Phe	
6720	6725	6730	
gtc agc gaa gtg cgc	gag cgc acc ggc ggg	cgt ggg gtg gat gtg	53653
Val Ser Glu Val Arg	Glu Arg Thr Gly Gly	Arg Gly Val Asp Val	
6735	6740	6745	
gtg ctc aac tcg ctg	gcg gga gag ctg ctc	ctg gcg ggc ctg tcc	53698
Val Leu Asn Ser Leu	Ala Gly Glu Leu Leu	Leu Ala Gly Leu Ser	
6750	6755	6760	
gtg ctg gcc ccg cac	ggc cgc ttc ctg gag	ctg ggg aag agg gac	53743
Val Leu Ala Pro His	Gly Arg Phe Leu Glu	Leu Gly Lys Arg Asp	
6765	6770	6775	
ctg tat gcg gac cag	cag gtg ggc ctc cga	acc ctg gcc cga ggg	53788
Leu Tyr Ala Asp Gln	Gln Val Gly Leu Arg	Thr Leu Ala Arg Gly	
6780	6785	6790	
cag act ttc gcc gcc	att gac ttc ggt ccc	cac cac ccg gac ttc	53833
Gln Thr Phe Ala Ala	Ile Asp Phe Gly Pro	His His Pro Asp Phe	
6795	6800	6805	
cga gcg gtg ctc gag	gag gtg gcc acg caa	ctc acc cag ggc cag	53878
Arg Ala Val Leu Glu	Glu Val Ala Thr Gln	Leu Thr Gln Gly Gln	
6810	6815	6820	
ctc gag cca ttg ccc	acc cgc ctc ttc ccc	gcg cgg cag gtg gcc	53923
Leu Glu Pro Leu Pro	Thr Arg Leu Phe Pro	Ala Arg Gln Val Ala	
6825	6830	6835	
gaa gcc ttc agc ttc	atg gcc cgc gcg ctg	cac atc ggt cgc gtc	53968
Glu Ala Phe Ser Phe	Met Ala Arg Ala Leu	His Ile Gly Arg Val	
6840	6845	6850	
gcc gtc tcc atg cag	ggg gcg acg gca ttg	ccc gcg tcc atg act	54013
Ala Val Ser Met Gln	Gly Ala Thr Ala Leu	Pro Ala Ser Met Thr	
6855	6860	6865	

cgg ggc tcc agg ccc Arg Gly Ser Arg Pro 6870	gca ccg gtg gcc gta Ala Pro Val Ala Val 6875	cct cct tgg gag gac Pro Pro Trp Glu Asp 6880	54058
ccg cgg ctg gcg ggc Pro Arg Leu Ala Gly 6885	ggc atc tcc tct gaa Gly Ile Ser Ser Glu 6890	gag ggt gcg gag gcc Glu Gly Ala Glu Ala 6895	54103
ttc ctg cgc gcg ttg Phe Leu Arg Ala Leu 6900	gag caa ggg gca ccg Glu Gln Gly Ala Pro 6905	cag ctc atc atc tcc Gln Leu Ile Ile Ser 6910	54148
ccc cag gac ttc agc Pro Gln Asp Phe Ser 6915	tcg ctg ctg cgc ggc Ser Leu Leu Arg Gly 6920	ctg ggc ggc agc cag Leu Gly Gly Ser Gln 6925	54193
ggc gtg cgc gaa aag Gly Val Arg Glu Lys 6930	gag cgc ctc gtc acc Glu Arg Leu Val Thr 6935	ggg cgc gcc gcc gcc Gly Arg Ala Ala Ala 6940	54238
gcc gag ccg cag gcc Ala Glu Pro Gln Ala 6945	ttg cca ccc tcc tcg Leu Pro Pro Ser Ser 6950	ctg gag cag ctc atc Leu Glu Gln Leu Ile 6955	54283
gag cag gtg tgg cgc Glu Gln Val Trp Arg 6960	aag cac ctg ggt gtg Lys His Leu Gly Val 6965	gag cgc gtg cag ccc Glu Arg Val Gln Pro 6970	54328
acg gac agc ttc ttc Thr Asp Ser Phe Phe 6975	cag ctt gga gga gac Gln Leu Gly Gly Asp 6980	tcg ctg ctg ggc atc Ser Leu Leu Gly Ile 6985	54373
caa gtg gcg gcg gat Gln Val Ala Ala Asp 6990	ctc cgc agg cac ctg Leu Arg Arg His Leu 6995	ggt gtg gag ctg ccc Gly Val Glu Leu Pro 7000	54418
acg gcc acc ctc ttc Thr Ala Thr Leu Phe 7005	agc cac ccc acc ctc Ser His Pro Thr Leu 7010	gcc gcg ctg gcc gcg Ala Ala Leu Ala Ala 7015	54463
gct ctg cgg gct cga Ala Leu Arg Ala Arg 7020	cag ggt gag gcc gcg Gln Gly Glu Ala Ala 7025	gct ccc act gcc ccc Ala Pro Thr Ala Pro 7030	54508
gcg cca gcg ctc gtg Ala Pro Ala Leu Val 7035	ccg gat cca gcc gcg Pro Asp Pro Ala Ala 7040	cgc ttc gag cct ttc Arg Phe Glu Pro Phe 7045	54553
ccc ctc acg gat gtg Pro Leu Thr Asp Val 7050	cag gaa gcc tac tgg Gln Glu Ala Tyr Trp 7055	gtg ggt cgc cgc tcg Val Gly Arg Arg Ser 7060	54598
gcc ttc gag ctc ggc Ala Phe Glu Leu Gly 7065	ggc gtc gcc gcc cat Gly Val Ala Ala His 7070	ggc tac ttc gaa atc Gly Tyr Phe Glu Ile 7075	54643
gac agc ccg ggg ctg Glu Ser Pro Gly Leu 7080	gag gtg gag cgc ttc Glu Val Glu Arg Phe 7085	atc caa tgc tgg cgc Ile Gln Cys Trp Arg 7090	54688
cag ctc ctg cag cgc Gln Leu Leu Gln Arg 54733	cac gac atg ctg cgc His Asp Met Leu Arg 54733	atg gtg gtg ctc ccg Met Val Val Leu Pro 54733	54733

7095	7100	7105
gat ggg cgg cag cag Asp Gly Arg Gln Gln 7110	gtg ctc gag cag gtg Val Leu Glu Gln Val 7115	ccg gag tac acg ccg Pro Glu Tyr Thr Pro 7120
54778		
gag gtg gtg gag ctg Glu Val Val Glu Leu 7125	cgg ggg ctc tcc ccc Arg Gly Leu Ser Pro 7130	cag gag gcc gag tcc Gln Glu Ala Glu Ser 7135
54823		
cgc cgg ctc cag ctg Arg Arg Leu Gln Leu 7140	cgt gag cgc atg gcc Arg Glu Arg Met Ala 7145	cac cag gtc ctg cgc His Gln Val Leu Arg 7150
54868		
agc gac cgc tgg ccg Ser Asp Arg Trp Pro 7155	ctc ttc gag ctg gtg Leu Phe Glu Leu Val 7160	ctc tgc cgg tac gag Leu Cys Arg Tyr Glu 7165
54913		
gga ggc gtc cgc atc Gly Gly Val Arg Ile 7170	cac atg agc atg gat His Met Ser Met Asp 7175	gcc ctg atg ctg gat Ala Leu Met Leu Asp 7180
54958		
gcg tgg agt tca gcg Ala Trp Ser Ser Ala 7185	gtg ctt cgg cag gac Val Leu Arg Gln Asp 7190	ttc gcc cag ctg tac Phe Ala Gln Leu Tyr 7195
55003		
cac gag ccg ggc cgg His Glu Pro Gly Arg 7200	ccg ctg gag ccg ctg Pro Leu Glu Pro Leu 7205	gcc atc acc ttc cgc Ala Ile Thr Phe Arg 7210
55048		
gac tac gtg ctg gcg Asp Tyr Val Leu Ala 7215	gag cgc cgg ctg cgc Glu Arg Arg Leu Arg 7220	gag ggc gag gcc cat Glu Gly Glu Ala His 7225
55093		
gag cgc gcc cgc gca Glu Arg Ala Arg Ala 7230	tac tgg tgg gct cgg Tyr Trp Trp Ala Arg 7235	ctg gac acg ctg ccg Leu Asp Thr Leu Pro 7240
55138		
cca ccg ccc gag ctg Pro Pro Pro Glu Leu 7245	ccc ctg gtg aag gaa Pro Leu Val Lys Glu 7250	ccc tcg cag ctg gag Pro Ser Gln Leu Glu 7255
55183		
cac gcg cgg ttc acc His Ala Arg Phe Thr 7260	cac cgc gag gct cgg His Arg Glu Ala Arg 7265	ctc gag cca cac cgc Leu Glu Pro His Arg 7270
55228		
tgg gcc cgg ctc cag Trp Ala Arg Leu Gln 7275	gag cgg gcg cgc gcc Glu Arg Ala Arg Ala 7280	cac ggc ctc acc ccc His Gly Leu Thr Pro 7285
55273		
tcg gcc gcc tgc atg Ser Ala Ala Cys Met 7290	gcc gcc ttc gcc gag Ala Ala Phe Ala Glu 7295	gtg ctt gcc cgc tgg Val Leu Ala Arg Trp 7300
55318		
agc cgt cac ccg cgc Ser Arg His Pro Arg 7305	ttc acc ctc aac ctc Phe Thr Leu Asn Leu 7310	acc ctc ttc cag cgc Thr Leu Phe Gln Arg 7315
55363		
ttg ccc ctg cac ccg Leu Pro Leu His Pro 7320	cag gtg gac gag ctg Gln Val Asp Glu Leu 7325	gtg ggc gac ttc acc Val Gly Asp Phe Thr 7330
55408		
tcc ctg gtc ctg ctg	gag gta gag gca cac	gcg gcg agc acc ttc
55453		

Ser Leu Val Leu Leu	Glu Val Glu Ala His	Ala Ala Ser Thr Phe	
7335	7340	7345	
gcc gag cgt gcc tcc	cgg ctc cag gca cag	cta tgg cgg gac ctg	55498
Ala Glu Arg Ala Ser	Arg Leu Gln Ala Gln	Leu Trp Arg Asp Leu	
7350	7355	7360	
gag cac ggc agc gtg	agc gcc gtg cag ctc	atc cgc gag ctc gtc	55543
Glu His Gly Ser Val	Ser Ala Val Gln Leu	Ile Arg Glu Leu Val	
7365	7370	7375	
cgc acc ggc cgc cgc	tcc ccg ggc gcc atc	atg ccc gtc gtc ttc	55588
Arg Thr Gly Arg Arg	Ser Pro Gly Ala Ile	Met Pro Val Val Phe	
7380	7385	7390	
acc agc atc ctc agc	ctg gat gcg cgg cgc	ggc ccc cag ggc agc	55633
Thr Ser Ile Leu Ser	Leu Asp Ala Arg Arg	Gly Pro Gln Gly Ser	
7395	7400	7405	
ctc tcc ttc ttc gag	gga gaa ctg gtg tac	agc atc agc cag act	55678
Leu Ser Phe Phe Glu	Gly Glu Leu Val Tyr	Ser Ile Ser Gln Thr	
7410	7415	7420	
ccc cag gtg tgg ctg	gac cac gga gtc cac	gag gag gag ggg gcg	55723
Pro Gln Val Trp Leu	Asp His Gly Val His	Glu Glu Glu Gly Ala	
7425	7430	7435	
ctc gtc ctg gcg tgg	gac tcg gtg gag gcg	ctc ttc cct ccg ggc	55768
Leu Val Leu Ala Trp	Asp Ser Val Glu Ala	Leu Phe Pro Pro Gly	
7440	7445	7450	
atg gtg gac gac atg	ttc cac gcc tac cag	cgg ctg ctg ggg gca	55813
Met Val Asp Asp Met	Phe His Ala Tyr Gln	Arg Leu Leu Gly Ala	
7455	7460	7465	
ctc gcc gag gag gag	caa gcg tgg gag ggc	gag ctg ccg gag ctg	55858
Leu Ala Glu Glu Glu	Gln Ala Trp Glu Gly	Glu Leu Pro Glu Leu	
7470	7475	7480	
ctg cct cct gcc cag	cgt gag ttg ctc gcg	cgc tac aac gcc acc	55903
Leu Pro Pro Ala Gln	Arg Glu Leu Leu Ala	Arg Tyr Asn Ala Thr	
7485	7490	7495	
cag gcg ccg cgg ccc	agc gga cgg ctg gag	gag ggc ttc ttc acc	55948
Gln Ala Pro Arg Pro	Ser Gly Arg Leu Glu	Glu Gly Phe Phe Thr	
7500	7505	7510	
cag gcg cgg ctc cac	ccc gag ctc ccc gcg	ctg ctc gca ccg gag	55993
Gln Ala Arg Leu His	Pro Glu Leu Pro Ala	Leu Leu Ala Pro Glu	
7515	7520	7525	
cgc acc ctg agc tac	ggc gag ctg gca agg	cga gcc cag gcg ctg	56038
Arg Thr Leu Ser Tyr	Gly Glu Leu Ala Arg	Arg Ala Gln Ala Leu	
7530	7535	7540	
gcc gcg cgc cta cgc	gag ctg gag gtg cag	cct cag gag ttg gtg	56083
Ala Ala Arg Leu Arg	Glu Leu Glu Val Gln	Pro Gln Glu Leu Val	
7545	7550	7555	
gcc att gcc atg cac	aag ggc tgg gag cag	gcc acc gcc gtg ctc	56128
Ala Ile Ala Met His	Lys Gly Trp Glu Gln	Ala Thr Ala Val Leu	
7560	7565	7570	

ggc gtg ctg cag gcg	gcc gcg gcc tac ctg	ccg ttg gat cca gag	56173
Gly Val Leu Gln Ala	Ala Ala Ala Tyr Leu	Pro Leu Asp Pro Glu	
7575	7580	7585	
cag ccg ccg ctc cgg	ctc cac cag ctc ctg	gag gag ggg ccc gcc	56218
Gln Pro Pro Leu Arg	Leu His Gln Leu Leu	Glu Glu Gly Pro Ala	
7590	7595	7600	
cgc gtg gtg ctc acc	cag tcc tcc ttg ctg	cac acc gtg ccc tgg	56263
Arg Val Val Leu Thr	Gln Ser Ser Leu Leu	His Thr Val Pro Trp	
7605	7610	7615	
ccg ccg ggt gtg cag	gtg atc gcc gtg gac	gaa ctc gag cct gca	56308
Pro Pro Gly Val Gln	Val Ile Ala Val Asp	Glu Leu Glu Pro Ala	
7620	7625	7630	
acc gag gcc ccc ccg	cta ccg cct cgc ggt	aca ccg gag cac ctc	56353
Thr Glu Ala Pro Pro	Leu Pro Pro Arg Gly	Thr Pro Glu His Leu	
7635	7640	7645	
gcc tac gtc atc tat	acc tcc ggc tcc acc	ggc aaa ccc aag ggt	56398
Ala Tyr Val Ile Tyr	Thr Ser Gly Ser Thr	Gly Lys Pro Lys Gly	
7650	7655	7660	
gtg gcc att gaa cac	cgg gcc gcg ctc aac	acc gtg gtg gac ctc	56443
Val Ala Ile Glu His	Arg Ala Ala Leu Asn	Thr Val Val Asp Leu	
7665	7670	7675	
aac acc cgc ttt ggc	gtc ggc ccg gag gat	cgg gtg ctg ggc ctg	56488
Asn Thr Arg Phe Gly	Val Gly Pro Glu Asp	Arg Val Leu Gly Leu	
7680	7685	7690	
tcc gcg ctc acc ttc	gac ctg tcg gtg tac	gac gtg ctg ggg ctg	56533
Ser Ala Leu Thr Phe	Asp Leu Ser Val Tyr	Asp Val Leu Gly Leu	
7695	7700	7705	
ctc ggc gcg ggc gga	gcg ctg gtg ctg ccc	gcg gcg gag gcg gag	56578
Leu Gly Ala Gly Gly	Ala Leu Val Leu Pro	Ala Ala Glu Ala Glu	
7710	7715	7720	
aag gat ccc gcc cac	tgg tgg gag cgg ctg	gtg gct ggg cgg gtc	56623
Lys Asp Pro Ala His	Trp Trp Glu Arg Leu	Val Ala Gly Arg Val	
7725	7730	7735	
acg gtg tgg aac tcc	acc ccg gcg ctg atg	ctg ctg ctg gtg gag	56668
Thr Val Trp Asn Ser	Thr Pro Ala Leu Met	Leu Leu Leu Val Glu	
7740	7745	7750	
tac gcc gag cag cgc	ggg ctg aag ctg ccc	gcg gcg ctg cgc ctg	56713
Tyr Ala Glu Gln Arg	Gly Leu Lys Leu Pro	Ala Ala Leu Arg Leu	
7755	7760	7765	
gtc atg ctc agc ggc	gac tgg atc ccc gtg	gcg cta ccg gat cgc	56758
Val Met Leu Ser Gly	Asp Trp Ile Pro Val	Ala Leu Pro Asp Arg	
7770	7775	7780	
atc cgc gca ctc ggc	agg gac gtg cag gtg	gtg agc ctg gga ggc	56803
Ile Arg Ala Leu Gly	Arg Asp Val Gln Val	Val Ser Leu Gly Gly	
7785	7790	7795	
gcc acc gag gct tcc	atc tgg tcc att gcc	tac ccc att ggc cag	56848
Ala Thr Glu Ala Ser	Ile Trp Ser Ile Ala	Tyr Pro Ile Gly Gln	
7800	7805	7810	

gtg gca ccg cag tgg Val Ala Pro Gln Trp 7815	aag agc att ccc tat Lys Ser Ile Pro Tyr 7820	ggc atg ccg ctg gcc Gly Met Pro Leu Ala 7825	56893
aat cag cga ttc cac Asn Gln Arg Phe His 7830	gtg ctg gac ggg cgg Val Leu Asp Gly Arg 7835	ctg gag gca cgg ccc Leu Glu Ala Arg Pro 7840	56938
tgg tgg gtg ccg ggc Trp Trp Val Pro Gly 7845	gag ctg tac att ggc Glu Leu Tyr Ile Gly 7850	ggc gag ggg ctg gca Gly Glu Gly Leu Ala 7855	56983
cgc gag tac tgg agg Arg Glu Tyr Trp Arg 7860	gat gag ccg ctc acc Asp Glu Pro Leu Thr 7865	gca acg cgc ttc atc Ala Thr Arg Phe Ile 7870	57028
cgt cac ccg cgc aca Arg His Pro Arg Thr 7875	ggc gag cgg ctg tac Gly Glu Arg Leu Tyr 7880	cgc acc ggg gac caa Arg Thr Gly Asp Gln 7885	57073
ggg cgc atg ctg ccc Gly Arg Met Leu Pro 7890	gaa ggg agc atc gag Glu Gly Ser Ile Glu 7895	ttc ctc ggc cgg gag Phe Leu Gly Arg Glu 7900	57118
gat ctg cag gtc aag Asp Leu Gln Val Lys 7905	gtg cag ggc ttc cgc Val Gln Gly Phe Arg 7910	gtg gag ctg ggg gag Val Glu Leu Gly Glu 7915	57163
att gaa gcg gcg ctg Ile Glu Ala Ala Leu 7920	gcc cag cac ccg gcg Ala Gln His Pro Ala 7925	cta tcg gcg agc gtg Leu Ser Ala Ser Val 7930	57208
gtg gtg gca cga gga Val Val Ala Arg Gly 7935	gag cct cgg gga gtg Glu Pro Arg Gly Val 7940	cgg cgg ctg gtg gcc Arg Arg Leu Val Ala 7945	57253
tac gcg gtt cct cgc Tyr Ala Val Pro Arg 7950	tcg ggt cag acg ccc Ser Gly Gln Thr Pro 7955	gcg gca ggc gag ctg Ala Ala Gly Glu Leu 7960	57298
cgc cgc tac ctg gcg Arg Arg Tyr Leu Ala 7965	gag cgg cta ccc gcg Glu Arg Leu Pro Ala 7970	tac atg gtg ccc tcc Tyr Met Val Pro Ser 7975	57343
gcc ttc gtc ctc ctg Ala Phe Val Leu Leu 7980	gag tcc ctg ccc cgc Glu Ser Leu Pro Arg 7985	agc cgc aac ggg aag Ser Arg Asn Gly Lys 7990	57388
atc gcc cga gat cag Ile Ala Arg Asp Gln 7995	ctc ccc gag ccg cag Leu Pro Glu Pro Gln 8000	cag aca cag gga ctg Gln Thr Gln Gly Leu 8005	57433
gct gca cag gcc gcg Ala Ala Gln Ala Ala 8010	gcg gcc gac ccg ctc Ala Ala Asp Pro Leu 8015	gtg gag cgg ctg gca Val Glu Arg Leu Ala 8020	57478
gcc ctg gtg aaa gag Ala Leu Val Lys Glu 8025	gcg ctg agg ctc gag Ala Leu Arg Leu Glu 8030	cgc gtg gag ccc cag Arg Val Glu Pro Gln 8035	57523
gac agc ctg ctg gac Asp Ser Leu Leu Asp	ctg ggg gcg gac tcg Leu Gly Ala Asp Ser	gtg gcg ctc atc cgc Val Ala Leu Ile Arg	57568

8040

8045

8050

ctc atc aac cgc ctg Leu Ile Asn Arg Leu 8055	gag gca gag ctg Glu Ala Glu Leu 8060	cag ttc cgc ccg cga ctg Phe Arg Pro Arg Leu 8065	57613
gcc gac atc tac gag Ala Asp Ile Tyr Glu 8070	aac ccc acc gtg Asn Pro Thr Val Gln 8075	ggg ctc gcc acg ctc Gly Leu Ala Thr Leu 8080	57658
cac caa gag aag aca His Gln Glu Lys Thr 8085	aag agc cag gga gag Lys Ser Gln Gly Glu 8090	gga ggc gct ccg cgg Gly Gly Ala Pro Arg 8095	57703
ctc acg gcg ccc cgc Leu Thr Ala Pro Arg 8100	tcc acg ctt ctg ccc Ser Thr Leu Leu Pro 8105	gcc gag gag tgg ggg Ala Glu Glu Trp Gly 8110	57748
cgc ttc aag gcc aac Arg Phe Lys Ala Asn 8115	cgc ccg ggc ctg cgc Arg Pro Gly Leu Arg 8120	cgc ttc ccc gat gga Arg Phe Pro Asp Gly 8125	57793
acg cct gag gtg gca Thr Pro Glu Val Ala 8130	ctg cca ggc agc ggc Leu Pro Gly Ser Gly 8135	ttg gcg ccg gcc ccc Leu Ala Pro Ala Pro 8140	57838
gag gag ctg aca gct Glu Glu Leu Thr Ala 8145	ctc gag cgc cgc cgc Leu Glu Arg Arg Arg 8150	agt gtg cgc acc tac Ser Val Arg Thr Tyr 8155	57883
tcc ctc gag ccc gtg Ser Leu Glu Pro Val 8160	agc cat gag cag ttg Ser His Glu Gln Leu 8165	gga cgg ctg ctg gcc Gly Arg Leu Leu Ala 8170	57928
ccc ttg cgc gag tgg Pro Leu Arg Glu Trp 8175	gag gtg cag ggc tcg Glu Val Gln Gly Ser 8180	cgg cgt tac ctc tac Arg Arg Tyr Leu Tyr 8185	57973
gcc tcc gcg ggc ggg Ala Ser Ala Gly Gly 8190	ctc tac ccg gtg cag Leu Tyr Pro Val Gln 8195	ctc tac ctc cac ctc Leu Tyr Leu His Leu 8200	58018
aag ccc ggc cgg gcg Lys Pro Gly Arg Ala 8205	cga ggg ctc gag ccc Arg Gly Leu Glu Pro 8210	ggc acc tgg tat tac Gly Thr Trp Tyr Tyr 8215	58063
gat ccc agc acc cat Asp Pro Ser Thr His 8220	cgg ctg gtg ctc ctg Arg Leu Val Leu Leu 8225	tcc gcc gga gcc ggg Ser Ala Gly Ala Gly 8230	58108
ctg gat cgc cgc atc Leu Asp Arg Arg Ile 8235	cat gat ccg cac cag His Asp Pro His Gln 8240	aac cag gcc atc ttc Asn Gln Ala Ile Phe 8245	58153
gac tcg gcc gcg ttc Asp Ser Ala Ala Phe 8250	tcc ctc ttc ctc atc Ser Leu Phe Leu Ile 8255	gcc cgc atg ggg gcc Ala Arg Met Gly Ala 8260	58198
gtg gag ccg gta tac Val Glu Pro Val Tyr 8265	gcc gaa cat gcc ctg Ala Glu His Ala Leu 8270	cac ttc gcc acg ctg His Phe Ala Thr Leu 8275	58243
gag gca ggg ctg atg	acc cag ctt ctc gat	ctg ggc gcg gct ccc	58288

Glu	Ala	Gly	Leu	Met	Thr	Gln	Leu	Leu	Asp	Leu	Gly	Ala	Ala	Pro	
8280					8285					8290					
agt	ggg	ctg	ggg	ctg	tgc	cat	atc	ggc	gac	ctg	gac	ttc	gcc	cag	58333
Ser	Gly	Leu	Gly	Leu	Cys	His	Ile	Gly	Asp	Leu	Asp	Phe	Ala	Gln	
8295					8300					8305					
gcg	cgg	ggg	ctc	ttc	cac	ttg	gag	gag	gag	cac	gtg	ctg	ctg	cac	58378
Ala	Arg	Gly	Leu	Phe	His	Leu	Glu	Glu	Glu	His	Val	Leu	Leu	His	
8310					8315					8320					
agc	ctg	gtg	ggc	gga	gta	ctc	ccc	act	cgg	ggg	cag	gag	gcg	gcc	58423
Ser	Leu	Val	Gly	Gly	Val	Leu	Pro	Thr	Arg	Gly	Gln	Glu	Ala	Ala	
8325					8330					8335					
tcg	gtg	ccc	gct	gag	gga	gga	acg	gag	gcg	cgg	caa	ctg	gcg	cag	58468
Ser	Val	Pro	Ala	Glu	Gly	Gly	Thr	Glu	Ala	Arg	Gln	Leu	Ala	Gln	
8340					8345					8350					
ctc	ctg	caa	cag	gtg	aag	acg	ctc	acg	ccc	gag	gcc	gcg	cga	gcg	58513
Leu	Leu	Gln	Gln	Val	Lys	Thr	Leu	Thr	Pro	Glu	Ala	Ala	Arg	Ala	
8355					8360					8365					
ctg	ctc	gag	gcc	cgc	cgc	ggc	agc	aag	ggg	aga	cca	cat	gag	taa	58558
Leu	Leu	Glu	Ala	Arg	Arg	Gly	Ser	Lys	Gly	Arg	Pro	His	Glu		
8370					8375										
gctgcacgag gagctggaga gcctggcgcc ggagcagcgg gagctactcg cgccctgat															58618
gaaagagcag ggccctggacg agggcgccatt gctgatgccc gtggagcgca agcccgaggg															58678
gctgccgctc tcgtccgccc agcagcgcat gtggttcctg caacagctca ggccctacgag															58738
ccccttctac aacgtgcacg cggcgctccg gctcacggga aagctgaagg tggagtgcct															58798
ggtgcacagc ctgaatgagt tcgtccggcg tcacgagccg ctgcgcacgg tgttcccctc															58858
cgccggaggt caaccctgc agcgcatacct cgcgcccgca ccggtgcgc tcgagcagcg															58918
cgacctgagc ggcgtgcccc cccaggagcg agaggccgag gtgtaccgcg cggtggagca															58978
cgcagcgctc gcctcctttg atctggagcg agagccgccc tgccgcttcc tcctggtgcg															59038
agtggagccg gaggaacacg tgctgggtgtt cgccacgcac cacatcgag cggtatggctg															59098
gtcgtctggc gtcttcgtgc aggagctgtg cgcgctgtac acggctgctg tacacgcgga															59158
gccaccggcg ctgccaccgc tcgggctgca gtacgcggac ttcgccgcgt gggagcggag															59218
ccggctcaag ggtggggcgcg agcgcgagct gctcgagtac tggcaggagc agctggcagg															59278
gttgccggac ctgagcaccc tgccgccaga gcggcccagg cccccgctgt ccaagcagga															59338
ggggccacg ttcgaattcg cctgccacc acaccaggtg caggccctgc gctcacttgc															59398
tcaagcgcg cgacctcgc tcttctccgt gctgctcgcc gccttccaat ggctgctggc															59458
cggctcagca ggccaggacg acgtggccct gggcatgccc atcgccaacc gcaaccgcaa															59518
ggcctcatcg gctgcttcgc cagtacctg gtgctgcgcg cgaagccctc															59578
gccttcacct catggctcgc ccaggtgtcc gagcagctcc acggcgccct															59638

cgagcaccag gaggtgcctt tcgaacgatt ggtggaggtg ctccagcccc ggcgccgcat 59698
ggatcggcac ccgctcttcc agatcttctt cgccatgcag cagcaccgcg tcgpgcgcg 59758
cgagctgcca ggcctgctgc tctcggagtt ccccttcgc agccgcgtgg cccgcttcga 59818
tctggagttc cacctctggg agtctgcccc gggcgctcgag ggcaccctca tctacgacgt 59878
ggacctgtat ggccaggagg gcgtagcgca actggtgcgg cgctacgtga gcctgctcga 59938
agccgtggcg gcggaccga cgcggacact gggggagctg gcgggtgaga cgctggcccc 59998
gcccgtacgg aagcaggttc tcgcccgtgc ggagtgtccg cccctccac gcccgagcac 60058
agcccctcgg aactggccg aggcacttct acagaccgcc gagcgcttcc ccacggccac 60118
ggtcagcttc gtccaggcac agggctcttg caccgctgg actctgccag agctggtgga 60178
gcgggccccg cgcctgcagg ccggcttgcg gcagtgggga ttgcggccgg gcgacagtct 60238
ggtgctggtg ctggggaggg aggaagagac ggtagaggca ctctgggcct gcgtgctggc 60298
gggagtggcg cccctggtgc taccggcgcc gcccgcgaga gcggaggcga gtcccgcgct 60358
ctcacgtctg cggcatgccc ggcagctgct ggggtgggcc cgggtcctca cccggcagga 60418
gatgctgccc gatcttgcca ggcagcttca ggtctcgccc accgcgcaca tcctgggggc 60478
cgtggaggag ctgagagcca cgggtggcga ggcgcccctg ccccagggc gcatggatga 60538
cgtggccctg ctcaacctca cctcgggcac caccggcaag gccagtgcg ccatgctcac 60598
gcaccgcaac ctgctggtgc gcctggaggc caccaacgtc gtctacgagt ccagccgct 60658
cgagcgcggc ctgctctggt tgcagctgca caacattggc gccctgtccg agtaccacct 60718
ccgcccgtc tcgcccggca tgcacacctt ccacgtccc accgaggagg tgctcgccga 60778
gcccctgca tggctggagt ggctcgagcg gtatggcatt gcacagacgt gggcgcccag 60838
cttcgcctac agccacctgc tggagcggct gcgcaagggt gaggaccggc gctggagcct 60898
gggaggcgtg cgcgtgctgc tcagcgccgg agagcagatc tccgcgcca tgggtggagga 60958
gctgatgcgg aggttggtc cctcgggagt gcgggaggat gccttcgtcg ctgcctgggg 61018
aatgaccgag acggcttcgg gggtagccta tgcgcggcgg cccggaacgc cgccccgcat 61078
gcacacctg gagcgcgca gcttgagcgg tccgcttcgg caccggcac ccgcttcgcc 61138
caccgcgtg cggctgatgg atgtaggag cccattgcc ggcaccgccc tgcgggtggt 61198
ggacgcgagc ggcgagctgc tctccagga gtgcgtgggg cgcattcaag tgcgcgcgca 61258
gatgatctct cccggctact acggggacc gaaggcctcg gcggcgctgc tcaccgccga 61318
cggctggctg gagacaggcg atctgggctt cctctccgag ggagcgctga ccattaccgg 61378
ccgcgccaag gatctggtca tcatccacgg caccaactt cctgctacg agatcgagtc 61438
gcgggtggag caggtggagg gagtggcgcc ctctcggcg gcggcgcgcg cgggtgcggat 61498
gctcgagggg agcagggagg agctggcggt cttcttcgtg cccacggagg gactcgctcc 61558

gcaacccccct gcctcactcc tgtcgcgcat ccgccagcag gtgctcgagc aggtgggggt 61618
 gcggatcgat catctcatcc cgctcgagcc tcaccagctc ccacgcacgg agggaggcaa 61678
 gctgcgccgc tcggagctga gggcgcgctt cgaagcaggt gagctgaggg cgccgcaacc 61738
 cgctccggtc ccgagccctt cgcgcccgct ggagcagctc atcgctccg tctgggcgga 61798
 ggtgctcgag caccaggaca tcgcgcccga ggccagcttc ttcgacctgg gcggcaactc 61858
 catcctcctg gtgcgcgtgg agcgggcact gcgcgcgcgg ctcggcctgg agctgacgct 61918
 catggacctg ttcgcctacc ccacggtgca ctcgctggcg gactacctgg agcctcgagc 61978
 cgcccagctg ccagcacagg cctccccctc cagcaggcc gagcgccggc gggggatgcg 62038
 cggcgcgggcg ctggagcaga gacgcactcg ccgccaggca caacgggacg aggactagac 62098
 acgc atg agc cac gac acg cag ctc gag gga gcc gta gcg gta gtg 62144
 Met Ser His Asp Thr Gln Leu Glu Gly Ala Val Ala Val Val
 8380 8385 8390
 gcc atg gcg ggg cgc ttc ccc cag gcc ccc acc ctg gag gac tac 62189
 Ala Met Ala Gly Arg Phe Pro Gln Ala Pro Thr Leu Glu Asp Tyr
 8395 8400 8405
 tgg cgc atc atc cac gag ggc gtg gac gcc ttt acc cag ctc acg 62234
 Trp Arg Ile Ile His Glu Gly Val Asp Ala Phe Thr Gln Leu Thr
 8410 8415 8420
 gac gaa cag ctc ctg gcc tcc gga gtg ggg ccc tcg ctc ctg cgc 62279
 Asp Glu Gln Leu Leu Ala Ser Gly Val Gly Pro Ser Leu Leu Arg
 8425 8430 8435
 caa ccc ggt tac gtg cgg cgc gcc gcc ctg ctg gac agg atg gac 62324
 Gln Pro Gly Tyr Val Arg Arg Ala Ala Leu Leu Asp Arg Met Asp
 8440 8445 8450
 cag ttc gac gcc gcg ttc ttt ggc ttc tcc ccg cga gag gcg gag 62369
 Gln Phe Asp Ala Ala Phe Phe Gly Phe Ser Pro Arg Glu Ala Glu
 8455 8460 8465
 gtg ctc gac cca cag cac cgg ctc ttc ctg gag tgc gcg cac gag 62414
 Val Leu Asp Pro Gln His Arg Leu Phe Leu Glu Cys Ala His Glu
 8470 8475 8480
 gcg ctg gag cgg gcc ggc cat ggc tcc gag cgc gca agg ggg cgc 62459
 Ala Leu Glu Arg Ala Gly His Gly Ser Glu Arg Ala Arg Gly Arg
 8485 8490 8495
 gtg ggg gtc ttc gcc agc gcc agc ctc aac agc tac tac ctg cat 62504
 Val Gly Val Phe Ala Ser Ala Ser Leu Asn Ser Tyr Tyr Leu His
 8500 8505 8510
 tcc ctg cac ggg aac gcc cgg ctg agg gag gtg ctg gga gac ttc 62549
 Ser Leu His Gly Asn Ala Arg Leu Arg Glu Val Leu Gly Asp Phe
 8515 8520 8525
 cag ctc gcc atc gcc aac gac aag gac ttc ctg ccc acc cgc gtc 62594
 Gln Leu Ala Ile Ala Asn Asp Lys Asp Phe Leu Pro Thr Arg Val
 8530 8535 8540

tcg tac Ser Tyr 8545	aag ctg ggc ctg cgc Lys Leu Gly Leu Arg 8550	ggg ccc agc gtg gcg Gly Pro Ser Val Ala 8555	gtg cag acg Val Gln Thr	62639
gca tgc Ala Cys 8560	tcc agc tcg ctg gtg Ser Ser Ser Leu Val 8565	gcc gtc cac ctc gcg Ala Val His Leu Ala 8570	tgc cag agc Cys Gln Ser	62684
ctg ctc Leu Leu 8575	aat ggt gag tgc gcc Asn Gly Glu Cys Ala 8580	ctg gcg ctc gcg gcc Leu Ala Leu Ala Gly 8585	ggc tcc tct Gly Ser Ser	62729
ctc tcc Leu Ser 8590	gtg ccc cag gcc cag Val Pro Gln Ala Gln 8595	ggc tac ctc tac cag Gly Tyr Leu Tyr Gln 8600	gag gga ggc Glu Gly Gly	62774
att gcc Ile Ala 8605	tca ccg gac ggg tac Ser Pro Asp Gly Tyr 8610	tgc cgc ccc ttc gac Cys Arg Pro Phe Asp 8615	gcg gcg gcg Ala Ala Ala	62819
gcg gcc Ala Gly 8620	acc aac cga ggc aac Thr Asn Arg Gly Asn 8625	ggc gtg ggg gtg gtg Gly Val Gly Val Val 8630	ctg ctc aag Leu Leu Lys	62864
ctg ctc Leu Leu 8635	gag gag gcg ctg gcg Glu Glu Ala Leu Ala 8640	gac gga gac acc atc Asp Gly Asp Thr Ile 8645	cac gcc gtc His Ala Val	62909
atc cgc Ile Arg 8650	ggc tcg gcg gtg aac Gly Ser Ala Val Asn 8655	aac gac ggg gca cac Asn Asp Gly Ala His 8660	aag atc ggc Lys Ile Gly	62954
tac acc Tyr Thr 8665	gcg cca agc gtg gag Ala Pro Ser Val Glu 8670	ggg cag gcc tcg gtc Gly Gln Ala Ser Val 8675	atc tcc gag Ile Ser Glu	62999
gcg ctg Ala Leu 8680	gag gtg gcg gcc gtc Glu Val Ala Ala Val 8685	tcc gcg gac agc atc Ser Ala Asp Ser Ile 8690	ggc tac gtg Gly Tyr Val	63044
gag gca Glu Ala 8695	cac ggc acc gcc acg His Gly Thr Ala Thr 8700	gcc ctg gga gat ccg Ala Leu Gly Asp Pro 8705	atc gag gtc Ile Glu Val	63089
cag gcc Gln Ala 8710	ctc acg cgg gcc ttc Leu Thr Arg Ala Phe 8715	cgc aag cac acg gag Arg Lys His Thr Glu 8720	cgc cgc ggc Arg Arg Gly	63134
tac tgc Tyr Cys 8725	gct ctc ggc tcg gtc Ala Leu Gly Ser Val 8730	aag gcc aac atc gga Lys Ala Asn Ile Gly 8735	cac ctg gat His Leu Asp	63179
gcg gca Ala Ala 8740	gcg ggc atc gcc agc Ala Gly Ile Ala Ser 8745	ctg atc aag gcc gtg Leu Ile Lys Ala Val 8750	ctg gtg ctg Leu Val Leu	63224
gag cgc Glu Arg 8755	cgc cag ctc ccg cca Arg Gln Leu Pro Pro 8760	tgc ccg cac ttc acc Cys Pro His Phe Thr 8765	tcc ccc aac Ser Pro Asn	63269
cca cgc Pro Arg 8770	atc gac ttc gag cgc Ile Asp Phe Glu Arg 8775	agc ccc ttc tac gtg Ser Pro Phe Tyr Val 8780	agc ggg cgg Ser Gly Arg	63314

ggg cag Gly Gln 8785	ccg tgg gag ccg gtg Pro Trp Glu Pro Val 8790	gac ggc ccg cgg cgc Asp Gly Pro Arg Arg 8795	gcg ggc gtc Ala Gly Val 8795	63359
agc tcc Ser Ser 8800	ttc ggt att ggt ggc Phe Gly Ile Gly Gly 8805	acc aac gtg cat gtg Thr Asn Val His Val 8810	gtg ctg gag Val Leu Glu 8810	63404
gag gcg Glu Ala 8815	ccg ccc cgg ccg gcg Pro Pro Arg Pro Ala 8820	ccg cgg cgg gca gcg Pro Arg Arg Ala Ala 8825	cgg ccc tac Arg Pro Tyr 8825	63449
cag ctc Gln Leu 8830	ctg ccg ctg tca gca Leu Pro Leu Ser Ala 8835	cgc acg gag cgc gcc Arg Thr Glu Arg Ala 8840	cgg gag gag Arg Glu Glu 8840	63494
gtc gag Val Glu 8845	gcc cgg ctg cgc gag Ala Arg Leu Arg Glu 8850	cac ctg cgc cag cag His Leu Arg Gln Gln 8855	ccc cag gag Pro Gln Glu 8855	63539
ccg ctg Pro Leu 8860	gag gag gtg gcc cac Glu Glu Val Ala His 8865	acg ctc cag gta ggc Thr Leu Gln Val Gly 8870	cgc cgc cac Arg Arg His 8870	63584
atg gcc Met Ala 8875	tgg cgc aca gcg ctg Trp Arg Thr Ala Leu 8880	gtg agc agc agc cct Val Ser Ser Ser Pro 8885	gcc cag gcg Ala Gln Ala 8885	63629
gta gag Val Glu 8890	ctg ctg gag agg cgc Leu Leu Glu Arg Arg 8895	cga ccc gag gcg ctg Arg Pro Glu Ala Leu 8900	ctg cgt ggc Leu Arg Gly 8900	63674
cag agt Gln Ser 8905	gcg gcc cag gcc cgc Ala Ala Gln Ala Arg 8910	tcc gtg gcc ttc ctt Ser Val Ala Phe Leu 8915	cta ccg ggc Leu Pro Gly 8915	63719
cag ggc Gln Gly 8920	tca cag tac gtg ggc Ser Gln Tyr Val Gly 8925	atg gga gcg gcg ctc Met Gly Ala Ala Leu 8930	cat gag tcc His Glu Ser 8930	63764
gag gcg Glu Ala 8935	ccc ttc cgt gaa cag Pro Phe Arg Glu Gln 8940	gtg gat ctc tgt gcc Val Asp Leu Cys Ala 8945	ggg aaa ctc Gly Lys Leu 8945	63809
cag ggc Gln Gly 8950	gtg ctc ggg ttg gac Val Leu Gly Leu Asp 8955	gtg cgc cgc ctg ctg Val Arg Arg Leu Leu 8960	tac cct ggc Tyr Pro Gly 8960	63854
ccg cgg Pro Arg 8965	gag cgg gaa tgg gcc Glu Arg Glu Trp Ala 8970	cag gag cgc ctg cgg Gln Glu Arg Leu Arg 8975	gag acg cgg Glu Thr Arg 8975	63899
gta tgc Val Cys 8980	cag ccg gtg ctc ttc Gln Pro Val Leu Phe 8985	acc gtg gag tac gcg Thr Val Glu Tyr Ala 8990	ctg gcc cga Leu Ala Arg 8990	63944
ctg ctg Leu Leu 8995	gag gcc tgg ggg ata Glu Ala Trp Gly Ile 9000	cgc ccc acc gcg ctg Arg Pro Thr Ala Leu 9005	ttg ggc cac Leu Gly His 9005	63989
agc ctg Ser Leu	ggc gaa tat gtg gcc Gly Glu Tyr Val Ala	gcg tgc ctc gcg ggc Ala Cys Leu Ala Gly	gtg ttc tcc Val Phe Ser	64034

9010	9015	9020	
ctg gaa gag gcg ctg gag gtg Leu Glu Glu Ala Leu Glu Val 9025	gtg gcc gca cgt gga Val Ala Ala Arg Gly 9030	cag ctc atg Gln Leu Met 9035	64079
ggc agc atg ccc ccg ggg ggc Gly Ser Met Pro Pro Gly Gly 9040	atg ctc gcc gtg Met Leu Ala Val 9045	ggg ctc cca gcc Gly Leu Pro Ala 9050	64124
cag gaa gtg gaa ccg ctc cta Gln Glu Val Glu Pro Leu Leu 9055	ccc gcg gca ctg gcc Pro Ala Ala Leu Ala 9060	ctg gcc gcg Leu Ala Ala 9065	64169
cac aac agc ccc cag gcc tgt His Asn Ser Pro Gln Ala Cys 9070	gtg gtg gcg ggc Val Val Ala Gly 9075	ccc acc gag ccc Pro Thr Glu Pro 9080	64214
ctg gcc gag ctg cgc gca cga Leu Ala Glu Leu Arg Ala Arg 9085	ctc gag cag cgg Leu Glu Gln Arg 9090	ggc atc gcg tgc Gly Ile Ala Cys 9095	64259
acg cct ctg gct gtc tct cac Thr Pro Leu Ala Val Ser His 9100	gcc ttc cac tcg ccg Ala Phe His Ser 9105	atg atg gag Met Met Glu 9110	64304
ccg gcg gtg gcc cca ttc gtc Pro Ala Val Ala Pro Phe Val 9115	gcc cgg ctg cgc cgg Ala Arg Leu Arg 9120	atg cat ctg Met His Leu 9125	64349
cgg cct ccg agc ctg ccc ttc Arg Pro Pro Ser Leu Pro Phe 9130	atc tcc aat gtg act Ile Ser Asn Val Thr 9135	ggc acc tgg Gly Thr Trp 9140	64394
ata gag gcc gag gag gcg acc Ile Glu Ala Glu Glu Ala Thr 9145	tca ccg gag tac tgg Ser Pro Glu Tyr Trp 9150	ggc cgg cac Gly Arg His 9155	64439
ctg ctg cag ccg gtg cgc ttt Leu Leu Gln Pro Val Arg Phe 9160	gcc cag ggg ctg gag Ala Gln Gly Leu Glu 9165	cgg ttg tgc Arg Leu Cys 9170	64484
gag ggt gtg cag ccg cgg ctg Glu Gly Val Gln Pro Arg Leu 9175	ctg ttg gaa gtg ggg Leu Leu Glu Val Gly 9180	ccg ggc cac Pro Gly His 9185	64529
aca ctg ggg cgg ctg gcg gcg Thr Leu Gly Arg Leu Ala Ala 9190	cgc caa agc tcc ggc Arg Gln Ser Ser Gly 9195	ccc gtg cag Pro Val Gln 9200	64574
gtg gtg tcc acg ctg ggc tcc Val Val Ser Thr Leu Gly Ser 9205	agc cgg gag gaa agc Ser Arg Glu Glu Ser 9210	tcg gag gtg Ser Glu Val 9215	64619
gag cgg ctg ctg acg gca atc Glu Arg Leu Leu Thr Ala Ile 9220	ggg cga ctg tgg gtg Gly Arg Leu Trp Val 9225	gag gga gcc Glu Gly Ala 9230	64664
gag gtg gac tgg gct ggg ctc Glu Val Asp Trp Ala Gly Leu 9235	cac cgc ggc gaa cgc His Arg Gly Glu Arg 9240	cgg agg cgc Arg Arg Arg 9245	64709
gcg gtg cta ccc acc tac cct 9250	ttt gag cac cag cgc 9255	tac tgg gtg 9260	64754

Ala Val 9250	Leu Pro Thr Tyr Pro 9255	Phe Glu His Gln Arg 9260	Tyr Trp Val	
gag aca Glu Thr 9265	gcc ccc cag cct tcc Ala Pro Gln Pro Ser 9270	cct ccc gag cgg cct Pro Pro Glu Arg Pro 9275	ggg aca cct Gly Thr Pro	64799
gcc gag Ala Glu 9280	agc cct gtc acc agc Ser Pro Val Thr Ser 9285	agc ttc tac gtg cca Ser Phe Tyr Val Pro 9290	ggt tgg agc Gly Trp Ser	64844
cgc gcg Arg Ala 9295	gct ctg ccc tcg gcg Ala Leu Pro Ser Ala 9300	gcc ccc tcg aga cgg Ala Pro Ser Arg Arg 9305	gtg ggc cgg Val Gly Arg	64889
ctg ctg Leu Leu 9310	ttg ttg gcg gag gcc Leu Leu Ala Glu Ala 9315	cat ggg tgg agc cag His Gly Trp Ser Gln 9320	ggg ctc gcg Gly Leu Ala	64934
gag cgg Glu Arg 9325	ctc cgg gcc cgg gga Leu Arg Ala Arg Gly 9330	cac tcc gtc acc ctc His Ser Val Thr Leu 9335	gtc gag cca Val Glu Pro	64979
ggt gag Gly Glu 9340	cgg ctc gag cgc ctg Arg Leu Glu Arg Leu 9345	acc ccg gag cac tgg Thr Pro Glu His Trp 9350	cgc ctg ccc Arg Leu Pro	65024
ccg ggc Pro Gly 9355	cgc cgc gag gac ttc Arg Arg Glu Asp Phe 9360	caa cgc ctg ctg gag Gln Arg Leu Leu Glu 9365	gac tcc gga Asp Ser Gly	65069
gag ctg Glu Leu 9370	ccc tcg agg gtg ctc Pro Ser Arg Val Leu 9375	cac ctc tgg ctg ctg His Leu Trp Leu Leu 9380	ggg agc agc Gly Ser Ser	65114
cag agc Gln Ser 9385	ccc cag gag cgc ggc Pro Gln Glu Arg Gly 9390	ttc tac acc ctg ctg Phe Tyr Thr Leu Leu 9395	gcg ctg gcc Ala Leu Ala	65159
cag gca Gln Ala 9400	ctg gga gcc cac ggc Leu Gly Ala His Gly 9405	acc cgg ccc gcg gtg Thr Arg Pro Ala Val 9410	gag ctc acg Glu Leu Thr	65204
gtg gtg Val Val 9415	acc gat cag ctc tac Thr Asp Gln Leu Tyr 9420	gcg gtg gca gag ggc Ala Val Ala Glu Gly 9425	gag ccg ata Glu Pro Ile	65249
cag ccg Gln Pro 9430	ctc aag gcg ctc ctg Leu Lys Ala Leu Leu 9435	cag ggc cct gct tcc Gln Gly Pro Ala Ser 9440	gtg ctt ccc Val Leu Pro	65294
cag gaa Gln Glu 9445	ctc ccc ggg tgc acg Leu Pro Gly Cys Thr 9450	tgc cgc ctc gtg gac Cys Arg Leu Val Asp 9455	atg gcc ctg Met Ala Leu	65339
ccg cca Pro Pro 9460	ggc ggc gtg gcg gag Gly Gly Val Ala Glu 9465	gag ctg ctg gag cgg Glu Leu Leu Glu Arg 9470	ctg gtg gcg Leu Val Ala	65384
gag ctc Glu Val 9475	gag tcc acc gcg tcc Glu Ser Thr Ala Ser 9480	gaa cgc tcg gtg gcg Glu Arg Ser Val Ala 9485	tac cgc ggc Tyr Arg Gly	65429

gcg gct Ala Ala 9490	cgc tgg gtg cgc gag Arg Trp Val Arg Glu 9495	ttc gtc ccc gtg ccg Phe Val Pro Val Pro 9500	ctg ccc ccg Leu Pro Pro	65474
ccc gag Pro Glu 9505	cct tcg cag ctc ccg Pro Ser Gln Leu Pro 9510	ctt cgg cag aga ggg Leu Arg Gln Arg Gly 9515	gtg tat ttg Val Tyr Leu	65519
att gtc Ile Val 9520	ggc ggc ctg ggg ggc Gly Gly Leu Gly Gly 9525	gta ggg ctc gcg ctg Val Gly Leu Ala Leu 9530	gca gag cac Ala Glu His	65564
ctg gcg Leu Ala 9535	cgg cgc gtg tcg gcg Arg Arg Val Ser Ala 9540	cgg ctg gtg ctc acg Arg Leu Val Leu Thr 9545	ggc cgc tcg Gly Arg Ser	65609
ccc acg Pro Thr 9550	cca ccg cgc gag agc Pro Pro Arg Glu Ser 9555	tgg agc gca tgg ctg Trp Ser Ala Trp Leu 9560	ggc act ccc Gly Thr Pro	65654
acg cgg Thr Arg 9565	ctg cgg ctc tca cag Leu Arg Leu Ser Gln 9570	gag ctc gag tgg ctg Glu Leu Glu Trp Leu 9575	cgc ggt gcc Arg Gly Ala	65699
gcc gag Ala Glu 9580	cag atc gag cac cag Gln Ile Glu His Gln 9585	cgg ccc ctg ctc tcc Arg Pro Leu His Ser 9590	ctg gcc gac Leu Ala Asp	65744
agt ccc Ser Pro 9595	ggc ctg gaa gag tcg Gly Leu Glu Glu Ser 9600	ctg cgc ctg ctg tgc Leu Arg Leu Leu Cys 9605	gcc agc tac Ala Ser Tyr	65789
ctc tac Leu Tyr 9610	cac ttt ctc ttc cct His Phe Leu Phe Pro 9615	ctt cag gcg cct ctg Leu Gln Ala Pro Leu 9620	aga ctt gga Arg Leu Gly	65834
gag ccg Glu Pro 9625	agg gcg atg cag gca Arg Ala Met Gln Ala 9630	ttg cgt gag cgc ctg Leu Arg Glu Arg Leu 9635	ggg ctg cag Gly Leu Gln	65879
ccc ggc Pro Gly 9640	ttc gag cgc ctc ttc Phe Glu Arg Leu Phe 9645	tcc ttc atg atc ggc Ser Phe Met Ile Gly 9650	acc ctg gag Thr Leu Glu	65924
cag gcg Gln Ala 9655	aag ctg att cga gtg Lys Leu Ile Arg Val 9660	gga gag ggc atg ctc Gly Glu Gly Met Leu 9665	gag ggt cag Glu Gly Gln	65969
gtg gag Val Glu 9670	cca gcc cac gtg ccc Pro Ala His Val Pro 9675	aca cct cga gca ctg Thr Pro Arg Ala Leu 9680	cac gag cga His Glu Arg	66014
ctg ctc Leu Leu 9685	gaa ggc tac ccg gag Glu Gly Tyr Pro Glu 9690	gtg tca ggc ctg ctg Val Ser Gly Leu Leu 9695	gag ctg ctt Glu Leu Leu	66059
gag cat Glu His 9700	tgt gtg cgc cac tac Cys Val Arg His Tyr 9705	cct gag gtg ctc cga Pro Glu Val Leu Arg 9710	ggg agc ctg Gly Ser Leu	66104
ccg gcc Pro Ala 9715	ctc agt gtg ctc tac Leu Ser Val Leu Tyr 9720	ccc tcg ggc cgc tcc Pro Ser Gly Arg Ser 9725	gag gag gac Glu Glu Asp	66149

tcc gca Ser Ala 9730	cga agc ggg gtc Arg Ser Gly Val 9735	gag Glu 9735	tgg agc agc atg Trp Ser Ser Met 9740	ggg Gly 9740	cag tgc gtg Gln Cys Val	66194
gcg ctg Ala Leu 9745	ctg cgg cat ttc Leu Arg His Phe 9750	ctc Leu 9750	gcc cgc cat Ala Arg His 9755	gcc gag Ala Glu 9755	cgc acg cag Arg Thr Gln	66239
ggg cgt Gly Arg 9760	acc ctg cgc atc Thr Leu Arg Ile 9765	ctc Leu 9765	gag gtg ggc ggc Glu Val Gly Gly 9770	ggc Gly 9770	agt ggc gtg Ser Gly Val	66284
gtg ctc Val Leu 9775	cag gcc ctg ctg Gln Ala Leu Leu 9780	ccg Pro 9780	ctg ctg cgg cac Leu Leu Arg His 9785	cac His 9785	ccg gtg gag Pro Val Glu	66329
tac cac Tyr His 9790	ttc acg gat att Phe Thr Asp Ile 9795	ggc Gly 9795	ccc tcc ttc gtc Pro Ser Phe Val 9800	cgc Arg 9800	gcc atg gag Ala Met Glu	66374
gga gtc Gly Val 9805	ggc cgc cgc cag Gly Arg Arg Gln 9810	ggc Gly 9810	ctc acc ttc ctc Leu Thr Phe Leu 9815	cac His 9815	acc tcg gtg Thr Ser Val	66419
ctg gac Leu Asp 9820	atc tcc cgc cct Ile Ser Arg Pro 9825	ccc Pro 9825	cca gag cag ggc Pro Glu Gln Gly 9830	cat His 9830	ccg cct ggc Pro Pro Gly	66464
tcg tat Ser Tyr 9835	gat ctg gtg att Asp Leu Val Ile 9840	gcc Ala 9840	ctc aac gtg gtg Leu Asn Val Val 9845	cac His 9845	gcc acc ccg Ala Thr Pro	66509
cgc gtt Arg Val 9850	cct cag tcc ctg Pro Gln Ser Leu 9855	gcc Ala 9855	cac ctc gag agc His Leu Glu Ser 9860	ctg Leu 9860	ctc gtc cca Leu Val Pro	66554
ggc ggc Gly Gly 9865	cac ctg tgc ctg His Leu Cys Leu 9870	gtg Val 9870	gag acg gtg aag Glu Thr Val Lys 9875	cag Gln 9875	cag ccc tgg Gln Pro Trp	66599
gtg gac Val Asp 9880	atg atc tgg ggt Met Ile Trp Gly 9885	ttg Leu 9885	gcc gag ggg tgg Ala Glu Gly Trp 9890	tgg Trp 9890	agc tac gag Ser Tyr Glu	66644
gac gag Asp Glu 9895	ctg cgc acc cgc Leu Arg Thr Arg 9900	tct Ser 9900	ccc ttg ctg gag Pro Leu Leu Glu 9905	gta Val 9905	ggc gac tgg Gly Asp Trp	66689
gag cac Glu His 9910	gcc ctg cgc gac Ala Leu Arg Asp 9915	gtg Val 9915	ggc ttt gcc gag Gly Phe Ala Glu 9920	gtg Val 9920	gag gtg ctg Glu Val Leu	66734
cca gcg Pro Ala 9925	gcc gtg gag cag Ala Val Glu Gln 9930	cgc Arg 9930	tca cgc tgg gac Ser Arg Trp Asp 9935	aac Asn 9935	gtg ctc ctc Val Leu Leu	66779
att gcc Ile Ala 9940	cag cgg ccc ggc Gln Arg Pro Gly 9945	gag Glu 9945	agc ggg ctt cag Ser Gly Leu Gln 9950	cca Pro 9950	gtg ggc ggg Val Gly Gly	66824
cgc gcg Arg Ala 9955	gcc atg cag gag Ala Met Gln Glu 9960	cgg Arg 9960	atc cgc cgg ctg Ile Arg Arg Leu 9965	cgg Arg 9965	gcc att gaa Ala Ile Glu	66869

9955	9960	9965	
gaa gcg gga ggc gag gtg ctg Glu Ala Gly Gly Glu Val Leu 9970	ccc ctg gtg gcg gac Pro Leu Val Ala Asp 9975	gtc acc gat Val Thr Asp 9980	66914
cgg gag cgc atg ggc gag gtg Arg Glu Arg Met Gly Glu Val 9985	ctc gcc gag gtg aag Leu Ala Glu Val Lys 9990	cgc cgc cat Arg Arg His 9995	66959
ggg gcg ctc cat ggg gtg atc Gly Ala Leu His Gly Val Ile 10000	cac gcc gcg ctg gtg His Ala Ala Leu Val 10005	ctg gag gat Leu Glu Asp 10010	67004
ggg ctg atg cag ctc aag acg Gly Leu Met Gln Leu Lys Thr 10015	cgc gag tgc gcc ggg Arg Glu Ser Ala Gly 10020	cgc gtg ctg Arg Val Leu 10025	67049
gcc tcc aag gtc gag ggc acc Ala Ser Lys Val Glu Gly Thr 10030	ctg gtg ctg gac Leu Val Leu Asp 10035	gaa gtc Glu Leu Leu Arg 10040	67094
gac gag cca ctc gac ttc ttc Asp Glu Pro Leu Asp Phe Phe 10045	gtc ctg tgc tcc tgc Val Leu Cys Ser Ser 10050	ctg agc gca Leu Ser Ala 10055	67139
ttg ctc ggc gcg ctg ggg cag Leu Leu Gly Ala Leu Gly Gln 10060	gcg gac tat gcc gcc Ala Asp Tyr Ala Ala 10065	gcc agt gcc Ala Ser Ala 10070	67184
ttc ctc gac gcg tat gcg cac Phe Leu Asp Ala Tyr Ala His 10075	tca cag cga ggg cgc Ser Gln Arg Gly Arg 10080	acg gac cgc Thr Asp Arg 10085	67229
cgc acc atc tcc atg gac tgg Arg Thr Ile Ser Met Asp Trp 10090	gat cgc tgg ctg gag Asp Arg Trp Leu Glu 10095	gtg ggc gcc Val Gly Ala 10100	67274
gcc atg cgc ctg ggg ctg ggg Ala Met Arg Leu Gly Leu Gly 10105	ctg gcg gcg gga gca Leu Ala Ala Gly Ala 10110	ctg ggg ctg Leu Gly Leu 10115	67319
cag cgc acg gca ccc gga gag Gln Arg Thr Ala Pro Gly Glu 10120	tac acg gtg cga tgg Tyr Thr Val Arg Trp 10125	cag gcc gag Gln Ala Glu 10130	67364
cgc tgc tgg tgg ctg gac gag Arg Cys Trp Trp Leu Asp Glu 10135	cac cgc ctg gaa gga His Arg Leu Glu Gly 10140	cgg gcc acg Arg Ala Thr 10145	67409
ctg ccg ggg gtg gcg tac ctg Leu Pro Gly Val Ala Tyr Leu 10150	gag ctg gtg cga gcc Glu Leu Val Arg Ala 10155	gcc ctg gtc Ala Leu Val 10160	67454
cag gaa ctg ggc gag gcc ccc Gln Glu Leu Gly Glu Ala Pro 10165	gtg gag ctg gag caa Val Glu Leu Glu Gln 10170	ctg gtg ctg Leu Val Leu 10175	67499
ctc tct ctg ctg gaa gcg ccc Leu Ser Leu Leu Glu Ala Pro 10180	gcg ggg gag gag gtg Ala Gly Glu Glu Val 10185	gag gtg cga Glu Val Arg 10190	67544
ttc cac ctg cga ccg gag gat	gag ggc tat gcg ctg	gag atc cgc	67589

Phe His 10195	Leu Arg 10200	Pro Glu Asp 10205	Glu Gly Tyr Ala Leu 10205	Glu Ile Arg	
agc cgg Ser Arg 10210	gct ggc ggc ctc gcg Ala Gly Gly Leu Ala 10215	aat ggg ggc tgg cgc Asn Gly Gly Trp Arg 10220	agc cat gcg Ser His Ala	67634	
atg ggc Met Gly 10225	cgg gtg agg atg ctg Arg Val Arg Met Leu 10230	cct cgc ggg agc gct Pro Arg Gly Ser Ala 10235	cga ccc acc Arg Pro Thr	67679	
cac ccg His Pro 10240	cta cgg gag ctg gag Leu Arg Glu Leu Glu 10245	gaa cgg ctc ggc ttg Glu Arg Leu Gly Leu 10250	acc cag gcg Thr Gln Ala	67724	
ccg agg Pro Arg 10255	gag cat gag cca gct Glu His Glu Pro Ala 10260	cag ggc ccc cgc gag Gln Gly Pro Arg Glu 10265	ctg gca cct Leu Ala Pro	67769	
gcc ctg Ala Leu 10270	ggc cgg cgc tgg tcc Gly Arg Arg Trp Ser 10275	tcg ctg agc tgg cac Ser Leu Ser Trp His 10280	cgc gcg tgg Arg Ala Trp	67814	
aag gga Lys Gly 10285	gaa gag ggc ctg gcc Glu Glu Gly Leu Ala 10290	ttg atc gag ctc ccc Leu Ile Glu Leu Pro 10295	gag gag ctg Glu Glu Leu	67859	
gcc gag Ala Glu 10300	gat ctg cgg caa tgg Asp Leu Arg Gln Trp 10305	ccc ctg cac cct gcc Pro Leu His Pro Ala 10310	ctg ctg gac Leu Leu Asp	67904	
gcg gcg Ala Ala 10315	acc ggg ttt gca ccg Thr Gly Phe Ala Pro 10320	ctg ccc ccc ggt gcc Leu Pro Pro Gly Ala 10325	tgg ctg ccg Trp Leu Pro	67949	
ctg tcc Leu Ser 10330	tat ggc aac acc cgc Tyr Gly Asn Thr Arg 10335	atc cac ggc ccg ctg Ile His Gly Pro Leu 10340	ccg cgg cag Pro Arg Gln	67994	
ctg tac Leu Tyr 10345	agc cac atc cgc cgg Ser His Ile Arg Arg 10350	ctg gag ccc gcc agt Leu Glu Pro Ala Ser 10355	gcc cag gcc Ala Gln Ala	68039	
gga gtg Gly Val 10360	gtg cgg ctg gag gtg Val Arg Leu Glu Val 10365	cgc ctc atg gac ggc Arg Leu Met Asp Gly 10370	gaa ggg cgg Glu Gly Arg	68084	
gag ctg Glu Leu 10375	ctg tgc gtg gag gag Leu Cys Val Glu Glu 10380	ttc gtg ctc cgc cgg Phe Val Leu Arg Arg 10385	gtg gag gtg Val Glu Val	68129	
gac tcg Asp Ser 10390	ttg gct cgg cca cag Leu Ala Arg Pro Gln 10395	ccc gct ggc aag gct Pro Ala Gly Lys Ala 10400	ccg cgc cag Pro Arg Gln	68174	
gca gtg Ala Val 10405	ctt cca cga cca ggc Leu Pro Arg Pro Gly 10410	gcg ctc gac tcg ctg Ala Leu Asp Ser Leu 10415	cga ctg caa Arg Leu Gln	68219	
ccc ctg Pro Leu 10420	gag cgc ctt ccg ccg Glu Arg Leu Pro Pro 10425	cag gag gga gag gtg Gln Glu Gly Glu Val 10430	gag gtg cag Glu Val Gln	68264	

gtg ctc Val Leu 10435	gcc gcg ggc ctc aat Ala Ala Gly Leu Asn 10440	ttc aag gac gcg ctg Phe Lys Asp Ala Leu 10445	ctg gcg ctg Leu Ala Leu	68309
ggg gcg Gly Ala 10450	ttg cca gtg gag ctg Leu Pro Val Glu Leu 10455	gcc aac ggc gcc cca Ala Asn Gly Ala Pro 10460	ctg gcg ctg Leu Ala Leu	68354
gga gtc Gly Val 10465	gag tgc gcg ggc atc Glu Cys Ala Gly Ile 10470	atc tcc gcg gta ggt Ile Ser Ala Val Gly 10475	cca gga gtg Pro Gly Val	68399
aga ggc Arg Gly 10480	ctg cga gtg ggc gag Leu Arg Val Gly Glu 10485	gcg gtg gtg gcc gcg Ala Val Val Ala Ala 10490	gca gca gga Ala Ala Gly	68444
gcc ttc Ala Phe 10495	gcc tcg cat gtg cga Ala Ser His Val Arg 10500	gtc ccc cag gag cag Val Pro Gln Glu Gln 10505	gtc ttc ccc Val Phe Pro	68489
aag cca Lys Pro 10510	gcg ggc ctg agc ttc Ala Gly Leu Ser Phe 10515	gag cag gcg gcg atg Glu Gln Ala Ala Met 10520	gtg ccc gtc Val Pro Val	68534
act ctg Thr Leu 10525	ttc acc gcc tgg tat Phe Thr Ala Trp Tyr 10530	gcc ctg gag gag ctg Ala Leu Glu Glu Leu 10535	gcg cgg ctg Ala Arg Leu	68579
cgc gcg Arg Ala 10540	gga gag cac gtc ctc Gly Glu His Val Leu 10545	atc cac gcc gcg gcc Ile His Ala Ala Ala 10550	acg ggg gtg Thr Gly Val	68624
gga ctg Gly Leu 10555	gcg ggg gtg aag ttg Ala Gly Val Lys Leu 10560	gcg ctg cgg cgg ggc Ala Leu Arg Arg Gly 10565	gcc acg gtg Ala Thr Val	68669
tac gcc Tyr Ala 10570	acc gcg ggc agc gag Thr Ala Gly Ser Glu 10575	ccg aag cgc gag ctg Pro Lys Arg Glu Leu 10580	ctg cgc tcc Leu Arg Ser	68714
ctg ggg Leu Gly 10585	gta acg ctg gcc atg Val Thr Leu Ala Met 10590	gac tcg cgc gcg ccc Asp Ser Arg Ala Pro 10595	ggg ttt gac Gly Phe Asp	68759
gac cag Asp Gln 10600	ctt ctg caa cac acc Leu Leu Gln His Thr 10605	cag ggg cga ggc gtg Gln Gly Arg Gly Val 10610	gac gtg gtg Asp Val Val	68804
ctc aac Leu Asn 10615	tcg ctg agc ggc gag Ser Leu Ser Gly Glu 10620	ttc ctc tcc cgc agc Phe Leu Ser Arg Ser 10625	ctg ggc gtg Leu Gly Val	68849
ctc gct Leu Ala 10630	cgc cac ggg cgc ttc Arg His Gly Arg Phe 10635	gtc gag ctg ggt gta Val Glu Leu Gly Val 10640	agg gac atg Arg Asp Met	68894
ctc tct Leu Ser 10645	ggc ggc acg ctg ccg Gly Gly Thr Leu Pro 10650	ctg gct ccc ttc gag Leu Ala Pro Phe Glu 10655	cgc ggg ctc Arg Gly Leu	68939
acc ttc Thr Phe 10660	ctc gcg gcg cag ata Leu Ala Ala Gln Ile 10665	gac ccg ggg atg aag Asp Pro Gly Met Lys 10670	ggc tac cgg Gly Tyr Arg	68984

gag ctg Glu Leu 10675	atg ggc gag gcc ctg Met Gly Glu Ala Leu 10680	cgg cag atc gag cgc Arg Gln Ile Glu Arg 10685	gga gag ctg Gly Glu Leu	69029
gag ccg Glu Pro 10690	ctg ccc tac acg gca Leu Pro Tyr Thr Ala 10695	tgg cca ctg gag cgc Trp Pro Leu Glu Arg 10700	gtg gcc gag Val Ala Glu	69074
gcg ctc Ala Leu 10705	cag ctc gtc tcc aag Gln Leu Val Ser Lys 10710	ggc agg cac gta ggc Gly Arg His Val Gly 10715	aag gtg gtg Lys Val Val	69119
ctc acc Leu Thr 10720	ccg gaa gag ccc ttg Pro Glu Glu Pro Leu 10725	tca cgg ccg cca gcg Ser Arg Pro Pro Ala 10730	ctc ccg agc Leu Pro Ser	69164
tcc tca Ser Ser 10735	agg cca gtg caa ccc Arg Pro Val Gln Pro 10740	cgg gca ggc gtg gcg Arg Ala Gly Val Ala 10745	att gtc ggg Ile Val Gly	69209
ctg cgc Leu Arg 10750	tcc gcc gag ggc tgc Ser Ala Glu Gly Cys 10755	gag gcc ttc gag cgc Glu Ala Phe Glu Arg 10760	atc ctc gcg Ile Leu Ala	69254
gcc ggc Ala Gly 10765	ctg ccg cag gtg gcc Leu Pro Gln Val Ala 10770	gtc tcc acc cgt gag Val Ser Thr Arg Glu 10775	ctg cgc gcg Leu Arg Ala	69299
cgc atg Arg Met 10780	gcg gag atc gag cgc Ala Glu Ile Glu Arg 10785	ctg cgc gtc tcc tcc Leu Arg Val Ser Ser 10790	tgg gag ccg Trp Glu Pro	69344
gcg ctt Ala Leu 10795	ccc tcg gtg ccc cgc Pro Ser Val Pro Arg 10800	ggc tcc ccc gcc agg Gly Ser Pro Ala Arg 10805	acc gaa cgc Thr Glu Arg	69389
tcc aag Ser Lys 10810	ccc tat gtt gct cct Pro Tyr Val Ala Pro 10815	cgc acc gag cgc gag Arg Thr Glu Arg Glu 10820	cgc gcc ctg Arg Ala Leu	69434
gcc gcg Ala Ala 10825	ctc tgg cag gac ctg Leu Trp Gln Asp Leu 10830	ctc ggc gtc tcc gag Leu Gly Val Ser Glu 10835	gtg ggg ttg Val Gly Leu	69479
gat gat Asp Asp 10840	gac ttc ttc gag ctg Asp Phe Phe Glu Leu 10845	aaa ggg gac tcg ctg Lys Gly Asp Ser Leu 10850	ctc gcc atc Leu Ala Ile	69524
cag ctc Gln Leu 10855	atg ggg ccg ttg cgc Met Gly Arg Leu Arg 10860	aag gag ctg gag gtg Lys Glu Leu Glu Val 10865	gaa ttg cca Glu Leu Pro	69569
ctc tcc Leu Ser 10870	tcc ttc ctc gcc cgc Ser Phe Leu Ala Arg 10875	cct acg ctg cgc aca Pro Thr Leu Arg Thr 10880	ctg ctc tcc Leu Leu Ser	69614
gcc atg Ala Met 10885	cca gcc ccg gca acg Pro Ala Pro Ala Thr 10890	gac gag cct gcc tcg Asp Glu Pro Ala Ser 10895	gcc acc ggg Ala Thr Gly	69659
gca atc Ala Ala 10895	cca gcc gcc gct ccc Pro Ala Ala Ala Pro 10900	cag ccg tca ccg gct Gln Pro Ser Pro Ala 10905	gtc tcc agc Val Ser Ser	69704

10900		10905		10910	
gaa ccc Glu Pro 10915	gct cgg aag tgg cgc Ala Arg Lys Trp Arg 10920	cac ctg gtg ccc atc His Leu Val Pro Ile 10925	caa cct cag Gln Pro Gln	69749	
ggc gag Gly Glu 10930	agc cct ccc ttc ttc Ser Pro Pro Phe Phe 10935	tgg gct gcg ccg ctg Trp Ala Ala Pro Leu 10940	atg ggc acc Met Gly Thr	69794	
gtg ttt Val Phe 10945	ccc tac ttc acg ctg Pro Tyr Phe Thr Leu 10950	gcc cgc tgc ctg gcg Ala Arg Cys Leu Ala 10955	ccc acc cac Pro Thr His	69839	
ccc ttc Pro Phe 10960	tat gct ctc cag ccg Tyr Ala Leu Gln Pro 10965	cca ggg ttg cag gag Pro Gly Leu Gln Glu 10970	ggg cag gag Gly Gln Glu	69884	
ccc atg Pro Met 10975	gac aag gtg gag gaa Asp Lys Val Glu Glu 10980	ctg gcg gcc ctg tac Leu Ala Ala Leu Tyr 10985	gtg cgc gag Val Arg Glu	69929	
ctg aga Leu Arg 10990	gag ctg cag ccg cga Glu Leu Gln Pro Arg 10995	ggg ccc tac cgc ctg Gly Pro Tyr Arg Leu 11000	gga ggc tgg Gly Gly Trp	69974	
tcc ttc Ser Phe 11005	ggc tgc gtg gtg gct Gly Cys Val Val Ala 11010	tat gag gtt gcc cta Tyr Glu Val Ala Leu 11015	cag ctc gag Gln Leu Glu	70019	
gcg gct Ala Ala 11020	gga gag cag gtg gcg Gly Glu Gln Val Ala 11025	ctg ctg agc ctg ctg Leu Leu Ser Leu Leu 11030	gac ttc cct Asp Phe Pro	70064	
gcc cct Ala Pro 11035	tcg ggc cag cga gcc Ser Gly Gln Arg Ala 11040	agt ctc tgg gcg acg Ser Leu Trp Ala Thr 11045	gcc cgc ttc Ala Arg Phe	70109	
ttc agc Phe Ser 11050	ggc tcg gtg ctg cgc Gly Ser Val Leu Arg 11055	ggc ctg gcg ccc tat Gly Leu Ala Pro Tyr 11060	acg ctc gac Thr Leu Asp	70154	
tat ctc Tyr Leu 11065	tac ctg gcg gct cgc Tyr Leu Ala Ala Arg 11070	gct cct ctg aat cca Ala Pro Leu Asn Pro 11075	gac gct gcc Asp Ala Ala	70199	
tcg cgc Ser Arg 11080	ctg agc agc tcc ctg Leu Ser Ser Ser Leu 11085	acc cga gga tgg gag Thr Arg Gly Trp Glu 11090	cgc ctg cgg Arg Leu Arg	70244	
aaa gga Lys Gly 11095	ggg ctg gtg cag gaa Gly Leu Val Gln Glu 11100	ctg ctg gac cat gcc Leu Leu Asp His Ala 11105	gcc atg gcc Ala Met Ala	70289	
cgg gtg Arg Val 11110	gtg cct gag gac tca Val Pro Glu Asp Ser 11115	cgg ctg ctg ctg ctg Arg Leu Leu Leu Leu 11120	cga gag cct Arg Glu Pro	70334	
ggc att Gly Ile 11125	gcc ccc atg ttg agg Ala Pro Met Leu Arg 11130	ctg acc cgg gct cac Leu Thr Arg Ala His 11135	cag cga gcc Gln Arg Ala	70379	
atg ctc	tcc tat cgg ccc cga	ggg cgc ctc cgc cag	cgc atc gtc	70424	

Met Leu	Ser Tyr Arg Pro Arg	Gly Arg Leu Arg Gln	Arg Ile Val
11140	11145	11150	
ctc ttc	cgc acc tcc gag cag	gcc tgg agc ttc gcg	cga gac ctg 70469
Leu Phe	Arg Thr Ser Glu Gln	Ala Trp Ser Phe Ala	Arg Asp Leu
11155	11160	11165	
ggc tgg	gga gcc ctc agc gcc	tgc ggc gtg gat gtg	cgc gag gcg 70514
Gly Trp	Gly Ala Leu Ser Ala	Ser Gly Val Asp Val	Arg Glu Ala
11170	11175	11180	
ccg gga	gac cat atg acg ctg	ctg cgg cca cct cac	gtg gag cgg 70559
Pro Gly	Asp His Met Thr Leu	Leu Arg Pro Pro His	Val Glu Arg
11185	11190	11195	
gtg gca	gag cag ctc cgg gcc	ctg ctc agc tcc gcg	cct ctc ccg 70604
Val Ala	Glu Gln Leu Arg Ala	Leu Leu Ser Ser Ala	Pro Leu Pro
11200	11205	11210	
act cgc	cgc tga gttcgggctg	cgcgccagag gctgcctcat	gcatgccctt 70656
Thr Arg	Arg		
11215			
tgccagccag	cccagagctgt ccacccactc	cagcgcctcg ctcagatcgt	gctttttcgac 70716
gatctgatcc	aaggtgacca cctgcgcttc	cacgctgttg cccagatgag	agggcacccg 70776
cggatcctct	ctccaataat tgcctcgcgg	cagcagcctc ttcgcctgga	agtcgatggc 70836
ctccactccc	gccgctacgg cgatgtccac	caaggcccag ggcaacttcg	agttcaacag 70896
ccacttccac	cagccccagt tctcgttgta	gcccttcagg tagtgcacca	tgtccccgcg 70956
cccctgggag	aacacgagga ggtcccgcag	gtccgcccgc ttggcgtaga	tcgcctgggc 71016
cacggctgcc	atggtgggggt tgttggcgaa	caagcctcca tccacatagc	cctgggtacac 71076
cggcgagacg	atcggcaggg agctggagcg	cagggccagg tctaccaggc	ggctctcctgg 71136
gttgacgaag	tcggcacccg gaaagtgtgtg	aaagactcgc ggcttcact	cccgtctgtt 71196
ctcgtcaggg	ttttcatcgt ccaacttgaa	ggcggggatc accaccttgc	gcttcaggtc 71256
cccagagcgtc	ttgtctccca ggatgtgggt	gagcgcacga gccatgttct	tgctgggcgt 71316
cagcgcgttc	atgccaatga tggcccccaa	tctgggtcctg agcgatgagt	ggaacagctc 71376
atgcgtgagc	cagagctgcc gcgcctgctc	gagccccacg gcaggctcgc	gagccgcccgc 71436
gatgatgagc	gcgctgatgg ctcccgcga	ggtgcccgcc accaggtcga	cctgatcaag 71496
gaagccgggc	cgcaactggc cgatgcgctt	caacaggcca atggacagca	gcgtccgcag 71556
tcctcctccg	tccaggctga ggatgcggta	cttgccattc cggcgagccc	cgctggctcc 71616
tggctgccc	tggatgatgg tcattttctcc	acctccggcg gtccgaaggc	gagatgccgc 71676
agtgggggc	accccatgtc aagcaggtgg	ccgtcaggcc gttcgaacct	ctgctcagct 71736
ccgcgcctct	ccgactcgc cactgaactc	acgctgcgcg ccagaggctg	tctcctgcgt 71796
ggcctgggg	tgccagccc agttctccac	ccactccagc accttgctca	gatcgtgcct 71856

cttgacgacc cgatccaagg tggcgatctg ggcctccaca ctgttgccca ggttgaaggg 71916
caccaccgga tcctccctcc agtagttgcc ccgcggaagc agcatctttg cctggaagtc 71976
gatggcctcc actcccgcct cgatggtggc ggccacgaag gccatgggct gcttcggggtc 72036
cagcagccac tttctccagc cccagttctc gttataaccg tccaggtagt ccaccgagtc 72096
gcctgtgccc agggagaaca cgaggatgtc ccgcacatcc gccgccttgg cgtagatggc 72156
ctggggccacg gctgacatgg tggggttgtt gggaacaag ccccatcca catagccctg 72216
atggaccggc gagacgatcg gcaggagct ggagcgcagg gccaggtcca ccaggtagtc 72276
atccaggttg acgaaagtgt caccgggaa gttgtggaag attcgcggcc tccaccgcg 72336
cctgtcctca tcagggtcgc catcgtccag ctggaaggcg gggatcacca ccttgcgctt 72396
caggtcccgg agcgtcttgt ctcccaggat gtgggtgagc gcgcgagcca tgttctcgct 72456
gggcatcagc gcgctctggc caaccaggc cccagctga ttggagagtg ggctatcgaa 72516
cagcccaggc gtgaaccaga tctgccgcgc ctgctcgagc cccacggcag gctcgcgagc 72576
cgccgcgatg atgagcgcgc tgatggctcc cgccagggtg cccgccacca ggtcgacctg 72636
atcaaggaag ccgggcccga actggtcgat gcgcttcaac agccaatgg acagcagcgt 72696
ccgcagtcct cctccgtcca ggctgaggat gcggtacttg tcattccggc gagccccgtc 72756
gcctcctgtt ttccgagtgc ccgaagtcac ttcgtcactc caggcagagg gctctcgggg 72816
ggcactctgc cctcacgaag gggaggtagc ccacggaagc gcttcgaggg cgaggcaggt 72876
actggcttgc cggagggatg ggggactcca tgcttgtcca ccggcaccgc gctctgtgca 72936
tgggactccg cgtcttccaa atcctgctcc accacctcca gcatggcgcg cagcaactcc 72996
ctgcttgccc gagtggcccc gagtcggagc gtctccatgg ccctcaggta cctgcccagc 73056
cctgccacgg gcattctctg cgtcactgcc gccgcggtac tgggcgtctg tccttgtatt 73116
tgtaccgcgt ccctgtagtc caggtttcgg ccctcccaa cgaaagtgtc atgcacagcc 73176
gggattcctc actcactgta catggcacct caggtgtctt ccaccgctt gcgctcgcat 73236
gttcatattc atcgtttgcc ctggagtgtg cgattgacct tgggctccgt gcgtagaagg 73296
ccgtgaaaat ccaagccgct cctccaaggc attgcgcgtc ctaacgatgg gaggggatcg 73356
gggaagtgca gtagccagaa aagcaacggg ctgccaggca gctatctaaa tgggcgttgc 73416
gtgcgtgcgc cccacctgcg ctcacagact tgggagcagc caacgcgtcag gtggagaacc 73476
atctagcgaa gtagatgctc ccacgggcaa tccccctgcc aaggcatgac ggaccaccaa 73536
tcccatgggt gctgggccac tccaagcgtg gcggcgccgg tgaaagacct tttcgcgcgc 73596
attgccaaac agttgcgcga aatcccaact ccctcctgac cggcgcgtcc cttcgtgtga 73656
cagatccggg ccaggatcgg acgaatcatc gtgtgctggc ttctcttcga gacgtgaagt 73716
tccacgcgac cgggcccgcc cggaggtgcc agccttactt tactgggtgtg attgtgatac 73776

11225	caa gcg	ccc gag acc tgg acg	gag cac tgg ttc gag	cac aac cag	75598
	Gln Ala	Pro Glu Thr Trp Thr	Glu His Trp Phe Glu	His Asn Gln	
11240		11245	11250		
acc gtc	tcc cgc gtc tac cag	gac aat gac gtg gcc	gtg tac ttc	75643	
Thr Val	Ser Arg Val Tyr Gln	Asp Asn Asp Val Ala	Val Tyr Phe		
11255	11260	11265			
gac tcg	gcc gtg aac cgg tcc	atc acc tgg ccc aac	acc tac ctg	75688	
Asp Ser	Ala Val Asn Arg Ser	Ile Thr Trp Pro Asn	Thr Tyr Leu		
11270	11275	11280			
ggc gac	gtg tgg cga tac acc	aag agg acc tac gga	cac ttc ggc	75733	
Gly Asp	Val Trp Arg Tyr Thr	Lys Arg Thr Tyr Gly	His Phe Gly		
11285	11290	11295			
acg gac	tcg cag ctc tac acc	atc ttc cac gca gga	aag tac agt	75778	
Thr Asp	Ser Gln Leu Tyr Thr	Ile Phe His Ala Gly	Lys Tyr Ser		
11300	11305	11310			
gga ggc	cac ccc tcc acg tac	ttc gac gcc agc cat	gac agg cgc	75823	
Gly Gly	His Pro Ser Thr Tyr	Phe Asp Ala Ser His	Asp Arg Arg		
11315	11320	11325			
aac gtg	att gat gtg ggt tcc	agt tcc ccc acg gcc	tgg act gcc	75868	
Asn Val	Ile Asp Val Gly Ser	Ser Ser Pro Thr Ala	Trp Thr Ala		
11330	11335	11340			
ggc acc	ggc aac gac ctg gac	atc gtc acc cat gag	gtg gcc cac	75913	
Gly Thr	Gly Asn Asp Leu Asp	Ile Val Thr His Glu	Val Ala His		
11345	11350	11355			
atc gtg	gaa ctc gcc agc aag	ggc gtg cac gac tcg	cct gcc ttc	75958	
Ile Val	Glu Leu Ala Ser Lys	Gly Val His Asp Ser	Pro Ala Phe		
11360	11365	11370			
ggc ctg	tgg cgc gac agc aag	tgg gcg gag atc ttc	aac tac gac	76003	
Gly Leu	Trp Arg Asp Ser Lys	Trp Ala Glu Ile Phe	Asn Tyr Asp		
11375	11380	11385			
gtc tac	ctg gcg ttg ggt cgc	acc agt gat gcc aat	cgc tgg tac	76048	
Val Tyr	Leu Ala Leu Gly Arg	Thr Ser Asp Ala Asn	Arg Trp Tyr		
11390	11395	11400			
aac ctg	atg gtc aac acg acg	gac agc ttc ccc cgc	gcc aac acc	76093	
Asn Leu	Met Val Asn Thr Thr	Asp Ser Phe Pro Arg	Ala Asn Thr		
11405	11410	11415			
cac tgg	ttc cgg gat tgg ttc	tat ccc atc tac aag	aac cac ggc	76138	
His Trp	Phe Arg Asp Trp Phe	Tyr Pro Ile Tyr Lys	Asn His Gly		
11420	11425	11430			
ggc tcc	tcg gtg ctc aac cgg	tac ttc gtc ctg ctc	gcg cag tac	76183	
Gly Ser	Ser Val Leu Asn Arg	Tyr Phe Val Leu Leu	Ala Gln Tyr		
11435	11440	11445			
ctg ccg	aag aac ggc agc aac	tac gcg cga agc ctg	aac tgg gga	76228	
Leu Pro	Lys Asn Gly Ser Asn	Tyr Ala Arg Ser Leu	Asn Trp Gly		
11450	11455	11460			
gag ttc	atc cac ttc tgg agt	ggc gcc gcg ggc gtc	aat ctc aag	76273	

Glu Phe 11465	Ile His Phe Trp Ser 11470	Gly Ala Ala Gly Val 11475	Asn Leu Lys
acc ctg Thr Leu 11480	gcc acg tcc gca ttc Ala Thr Ser Ala Phe 11485	gga tgg ccc gcc gag Gly Trp Pro Ala Glu 11490	tgg gaa gcg Trp Glu Ala 76318
cag ttc Gln Phe 11495	acc cag gct cag cgt Thr Gln Ala Gln Arg 11500	gac ttc ccc ttc acc Asp Phe Pro Phe Thr 11505	tac acc ccc Tyr Thr Pro 76363
ccg gcc Pro Gly 11510	ccc acg gct gtc acc Pro Thr Ala Val Thr 11515	gtg ttc cag gac cag Val Phe Gln Asp Gln 11520	aac tac ggt Asn Tyr Gly 76408
ggc tat Gly Tyr 11525	gga atg gca ctg ccc Gly Met Ala Leu Pro 11530	gtt ggc cgt tac acc Val Gly Arg Tyr Thr 11535	ctg tcc gcc Leu Ser Ala 76453
ctg cat Leu His 11540	gcc tgg ggc gtg cgc Ala Trp Gly Val Arg 11545	aat gac gac atc acc Asn Asp Asp Ile Thr 11550	tcc ctg aag Ser Leu Lys 76498
gtc gcc Val Ala 11555	agc ggc tac aag gtc Ser Gly Tyr Lys Val 11560	acc ctg tat gaa cat Thr Leu Tyr Glu His 11565	gac aac ttc Asp Asn Phe 76543
agc ggt Ser Gly 11570	gcc agc ctc acc aag Ala Ser Leu Thr Lys 11575	acg gcg gac gat gcc Thr Ala Asp Asp Ala 11580	tcg ctc gtc Ser Leu Val 76588
gac gac Asp Asp 11585	agc tgg aac gac aaa Ser Trp Asn Asp Lys 11590	atg tcc tcg ctc atc Met Ser Ser Leu Ile 11595	gtg agc gcg Val Ser Ala 76633
tcg ggg Ser Gly 11600	atg tag gtcagcgccg Met	ctcgggctcc ggggcgtgcc gggcttcctt	76685
cgcgagccct	ccccggcgcg	ttcgcgccag ggggaatggcg ggccgcccac	ggattggaga 76745
atgggtcgcc	cccaacgggc	ccaccgcggt acacagtgcg c atg ttc gac gcc	Met Phe Asp Ala 76798
			11605
act cgc tac ctg gag Thr Arg Tyr Leu Glu 11610	cgt atc ggt gtc tcg Arg Ile Gly Val Ser 11615	gcg ggt acg ccg ctc Ala Gly Thr Pro Leu 11620	76843
gcc gcg ctc cat cgg Ala Ala Leu His Arg 11625	gcc cac ctc gag gcc Ala His Leu Glu Ala 11630	gtg ccg ttc gag aac Val Pro Phe Glu Asn 11635	76888
ctc gac atc cac ctg Leu Asp Ile His Leu 11640	aag cgg ccc atc cgg Lys Arg Pro Ile Arg 11645	ctc gat gag gcg gcc Leu Asp Glu Ala Ala 11650	76933
ttc ttc gag aag gtc Phe Glu Lys Val 11655	gtc ctc cag cgg cgc Val Leu Gln Arg Arg 11660	ggg ggg ttc tgc tac Gly Gly Phe Cys Tyr 11665	76978
gag gtc aat ggc ctc Glu Met Asn Gly Leu 11670	ttc gcc cgc ctg ctc Phe Ala Arg Leu Leu 11675	cgg act ctc ggt tac Arg Thr Leu Gly Tyr 11680	77023

'

ccg gcc ccc Pro Ala Pro 11880	aag aag tac gtg gat Lys Lys Tyr Val Asp 11885	tac acc gtg aag agc Tyr Thr Val Lys Ser 11890	ggt gac Gly Asp 77885
acg ttc tcg Thr Phe Ser 11895	ggc atc gcg gcg aag Gly Ile Ala Ala Lys 11900	cac gag ctc tcg ctg His Glu Leu Ser Leu 11905	gcc gca Ala Ala 77930
ttg aag aag Leu Lys Lys 11910	ctg aac ccg cag gtc Leu Asn Pro Gln Val 11915	gag aac atc aac gtc Glu Asn Ile Asn Val 11920	atc cac Ile His 77975
ccg ggc cag Pro Gly Gln 11925	aag ctg cac gtg aag Lys Leu His Val Lys 11930	gtc gtg aat gcg ccc Val Val Asn Ala Pro 11935	gcg cct Ala Pro 78020
gag ccg aag Glu Pro Lys 11940	ccg aca ccg aag ccg Pro Thr Pro Lys Pro 11945	ccg acc aac aaa aac Pro Thr Asn Lys Asn 11950	acc agc Thr Ser 78065
gtg ctc ccc Val Leu Pro 11955	aag ggc atc ccg aac Lys Gly Ile Pro Asn 11960	acc gag ggc atg tct Thr Glu Gly Met Ser 11965	cag gcg Gln Ala 78110
aag gag tac Lys Glu Tyr 11970	gcg ctc tac tcg aag Ala Leu Tyr Ser Lys 11975	tac gtg gag aag tac Tyr Val Glu Lys Tyr 11980	ggc gac Gly Asp 78155
gcc aag gcg Ala Lys Ala 11985	aag cag gac ctc gcg Lys Gln Asp Leu Ala 11990	gcc ggc aag cgc gtc Ala Gly Lys Arg Val 11995	atc gtc Ile Val 78200
ggc ctg agg Gly Leu Arg 12000	gtg aac acg ccc ttc Val Asn Thr Pro Phe 12005	acc cgt gat cgc ccc Thr Arg Asp Arg Pro 12010	tcc ggc Ser Gly 78245
ggc tcc tac Gly Ser Tyr 12015	gac gat cgg ctc gtc Asp Asp Arg Leu Val 12020	gtc atg tgg aag gac Val Met Trp Lys Asp 12025	agc agc Ser Ser 78290
ggc aag ccg Gly Lys Pro 12030	cac gtc cag gag ttc His Val Gln Glu Phe 12035	aag gcc aac acc gag Lys Ala Asn Thr Glu 12040	ccc aac Pro Asn 78335
cgc cgc tgg Arg Arg Trp 12045	gcc gat gac ccg agc Ala Asp Asp Pro Ser 12050	cag agc acc aag ccg Gln Ser Thr Lys Pro 12055	gtc ggc Val Gly 78380
cgc ctc gtc Arg Leu Val 12060	ggc aac aag acc tac Gly Asn Lys Thr Tyr 12065	cac tac cgg aag agc His Tyr Arg Lys Ser 12070	ttc aac Phe Asn 78425
ggc aac ttc Gly Asn Phe 12075	ggt ggc aac atc ctc Gly Gly Asn Ile Leu 12080	gcg ccg gac ctg cgc Ala Pro Asp Leu Arg 12085	tac ggg Tyr Gly 78470
aac ccg acc Asn Pro Thr 12090	gtg cgg cgc gac acc Val Arg Arg Asp Thr 12095	aac cgc gat cac cgg Asn Arg Asp His Arg 12100	atc aac Ile Asn 78515
tcg aag gac Ser Lys Asp	gat gtg ttc agc ggc Asp Val Phe Ser Gly	gac tgg ggc ggc cag Asp Trp Gly Gly Gln	ggc tac Gly Tyr 78560

12105	12110	12115	
tac ttc cat cgc ggc gga acg acc	gac acg tac tcc gcc	ggc tgc	78605
Tyr Phe His Arg Gly Gly Thr Thr	Asp Thr Tyr Ser Ala	Gly Cys	
12120	12125	12130	
cag acg atg gac cag ggc cgc ttc	aac aac ttc tgg gcc	gcg ctc	78650
Gln Thr Met Asp Gln Gly Arg Phe	Asn Asn Phe Trp Ala	Ala Leu	
12135	12140	12145	
ggc tct cag agc gag ttc agc tac	gtc ctg gcc cag gtc	ggc tga	78695
Gly Ser Gln Ser Glu Phe Ser Tyr	Val Leu Ala Gln Val	Gly	
12150	12155	12160	
cgtggcgtag gggaccact ccatccgga	gacccggaca cccgccgcgc	gagcaggagt	78755
cgtaagacct cgccgtgccc cgctcgaacg	ggctccagcc ttcgtgagac	ttcccttggc	78815
cacacgccag gaggtggtcc atggagaagg	agctggagca gttccgtcag	gaagcgcagc	78875
gattgaaggc cgggcggcgc ggtggctcgt	tgccattccc cgaggcgctg	cgcgcccttcg	78935
cggtgcgcta cgccgagcac accatggcga	agggtgggac ggtgacggac	gcggcgcagc	78995
aactgggcgt gtccgggccc acgctctacg	agtggcgcaa ggggcgcccc	gcgggacacc	79055
gccgtccgaa gcccgccgag aaggcgcg	ctctggtgcc ggtgcacgtc	agcgagcgtc	79115
ctgccgcacc tgtcaggtag tc atg acg	cgc gtc gag cag ctg cga gac		79164
	Met Thr Arg Val Glu Gln Leu Arg Asp		
	12165	12170	
ttg gag acg gcc aag cag ggg gcg gcg	ctg ctg gag gcg gaa aac		79209
Leu Glu Thr Ala Lys Gln Gly Ala Ala	Leu Glu Ala Glu Asn		
12175	12180	12185	
gcc cgg ctg cac cag cgc ctt gag gcg	ctc gtc gcg gag aat gcc		79254
Ala Arg Leu His Gln Arg Leu Glu Ala	Leu Val Ala Glu Asn Ala		
12190	12195	12200	
cgg ctc aag ggc gag gac gcc cag gcg	cac ctg cac cac gcg gat		79299
Arg Leu Lys Gly Glu Asp Ala Gln Ala	His Leu His His Ala Asp		
12205	12210	12215	
gag acg cat tgg ctg ctg ctg gac aag	ggg ccg ggc acc aag tgg		79344
Glu Thr His Trp Leu Leu Leu Asp Lys	Gly Pro Gly Thr Lys Trp		
12220	12225	12230	
tac gcg tgg acg gtg gcc tcg ccc gac	acc gtc ttc cat cgc atc		79389
Tyr Ala Trp Thr Val Ala Ser Pro Asp	Thr Val Phe His Arg Ile		
12235	12240	12245	
tac ccc agc cgc tcc ggg gcg acg gcg	cgc acc gtg ctg ggc gac		79434
Tyr Pro Ser Arg Ser Gly Ala Thr Ala	Arg Thr Val Leu Gly Asp		
12250	12255	12260	
tac gag ggc gtc gtg ctg gtg gat ggg	tat gcg gcg tac cag acg		79479
Tyr Glu Gly Val Val Leu Val Asp Gly	Tyr Ala Ala Tyr Gln Thr		
12265	12270	12275	
gcg acg aag tcc agc gcg gac ggg ccc	tgt cca gcc acc ctg gcc		79524
Ala Thr Lys Ser Ser Ala Asp Gly Pro	Cys Pro Ala Thr Leu Ala		
12280	12285	12290	

ttc tgc tgg gcc cac	gtg agg aga aaa ttc	ttc gaa gcc cag cag	79569
Phe Cys Trp Ala His	Val Arg Arg Lys Phe	Phe Glu Ala Gln Gln	
12295	12300	12305	
ttt gcc cct gcc tgc	aaa gag gtg ctg gac	ctc atc ggc gag ctg	79614
Phe Ala Pro Ala Cys	Lys Glu Val Leu Asp	Leu Ile Gly Glu Leu	
12310	12315	12320	
tac gcc gtc gag gcg	gac ctg ccg ggc tgg	tac gcg ctg gag ggc	79659
Tyr Ala Val Glu Ala	Asp Leu Pro Gly Trp	Tyr Ala Leu Glu Gly	
12325	12330	12335	
gag gag cgt cag gcc	gcg ctc gcg cac cgg	ctc gcc gtg cgt cag	79704
Glu Glu Arg Gln Ala	Ala Leu Ala His Arg	Leu Ala Val Arg Gln	
12340	12345	12350	
cag aag tcc gcg ccg	ctg acc gag cgc atc	cgc gac tgg gcg cac	79749
Gln Lys Ser Ala Pro	Leu Thr Glu Arg Ile	Arg Asp Trp Ala His	
12355	12360	12365	
gcg cag cgc gca ctg	ccc ggc agt gcc ttt	cgc aag gcc ctc gag	79794
Ala Gln Arg Ala Leu	Pro Gly Ser Ala Phe	Arg Lys Ala Leu Glu	
12370	12375	12380	
tac atg ctc aac ctc	tgg gcg ggg ctg acg	gtc ttc ctc acc cag	79839
Tyr Met Leu Asn Leu	Trp Ala Gly Leu Thr	Val Phe Leu Thr Gln	
12385	12390	12395	
ccg cag gtg ccg ctg	gac aat aac cac gtc	gag cgc cag ctg cgt	79884
Pro Gln Val Pro Leu	Asp Asn Asn His Val	Glu Arg Gln Leu Arg	
12400	12405	12410	
gac atg gtg ata ggc	agg aag aac ctc aag	aca gcg cgg ctg cga	79929
Asp Met Val Ile Gly	Arg Lys Asn Leu Lys	Thr Ala Arg Leu Arg	
12415	12420	12425	
gga gag gcc ccg ggc	cac tac ctg cgg cgc	cct gcg ctc gcc gcc	79974
Gly Glu Ala Pro Gly	His Tyr Leu Arg Arg	Pro Ala Leu Ala Ala	
12430	12435	12440	
atc gag aat ccg ggc	acc gtg acg ctg ccg	aag gcc cag agc tga	80019
Ile Glu Asn Pro Gly	Thr Val Thr Leu Pro	Lys Ala Gln Ser	
12445	12450		
ccaccgctcc cgctgtgcgg	cccggggcggg tcagctctta	gttcgggctc agcgcggaagg	80079
cggacacccg gccatccggg	atgagcgcca cccggtagcg	cagcatcgtg gagccggcct	80139
gcaccaggta cgtgtagacg	cggtcatcgc cccgcatctc	ccgcttcgcc aggacgagcg	80199
gccccgacgg tcccaggctc	tgtagcagcg actggtaata	cggcgcggcc tccgtgaaga	80259
accacccggg gacgtacgcg	aactcggcag ggttcagcgc	tccggcccgt gcctgctcga	80319
gcagcgtcgc gagccggggc	gtgacctccg gttcgaggtc	cggaatgggg gccagcggag	80379
gcacggcgag cgcgggggtt	acgatggcgg cgatgccatt	ggcgatggcc gcggggttcg	80439
gctccgagag gttgagcagc	acgatgatgg acagggcgtc	tccgggtgaag cggtagtagg	80499
cggctcttgaa gccctgccat	gcaccggcgt gctggtgcac	gagctggccg ttgcgctcct	80559

cgatgaacca tccgaagcca tacggatagg aagctccgct ggagagcttc acgggactca 80619
ggatgtcccg ccagctcccg ttcgtcagga tggcgcggtc atccacggcg tcatcccacg 80679
ccagcatgtc tttcagggag aagtagagcg agccatccgc cgtggtgttg agggaggggg 80739
agacccactc ctggtttctc accaccccggt cgagcaggcg gtaaccggcc gcgcgattgg 80799
gaacaatgtc agcctcactg atgcctcgcg ccgtcttcat gcgggtgggc ttgaagacct 80859
gctcgcccag gacgttgacg taggaggtgc ccgcgacgcg gttcaccagg agggccagca 80919
ggacgtagcc ggcgttgctg tagttccatc gcaggccggc ggggaagtcc aggggcagcg 80979
tatagaggaa ctgcgcgaac tcctcgtcg ttagtgcctt gcgcaggtcg agcagcccat 81039
ccgagaggcc ggaggtgtgc gtcagcagggt ggcgaatcgt gatggggggc caggtcgagg 81099
gcgcgtcggg gaagtacttc gtgatgctgt cggacagcga gatgcggccc gcctccacct 81159
ggagcatcac cgccatggcg gtgaactgct tgcccaggga gccggactgg aagagggtgt 81219
ccggggtgac gggaacctga tgctccagg tgcagaaacc atacccttc gccacgatga 81279
ccctgccgtg gctcacgacg gccaccgcga cgcccgccac attcatccgt ttctgctccg 81339
cgcgcacgaa ggcacgatg cgctcacgga gcgattcatg ggctgacgag ggacggcccg 81399
cgagcagacc cacgaggaac accatcgcca gggaaaagga ctgcaacgaa cctcgcatgg 81459
ggaactccag aggtcccggg cacggattgc ccacggacc gcgtagggga acaccgtcct 81519
tcctccgtcg taaaggcttc gggcggtga gaactacgat agttgtcgta tgaggggttc 81579
tcctggactc tcatccgctt ccggtttagc cgccgctccc gagcgaagcc tcacgcttgc 81639
gtcgttcccg ctctcggcg tgctggcgtg agcgggcccgt cacctcggca atgagggggt 81699
acggcgagga gttaccagga gtcagccctc tttcccgggc tcgagcgggg cgagcggaag 81759
ccgcccagac tcctgcgcgc gcgcgtgcac cgcgcgacc gccgcgtcca tcgcatcccg 81819
gacggccgcg ctctcgaaca cctcaagcc ggccagcgtc gcgccaccag ggctcgtgac 81879
ctgcttgccg aacgcgcgcg gggactcctg gctgcgatgc agcaactcca ggcttccgcg 81939
catgctctcc accgtcgcca ggggtggcctg atcacgcggc agccccaggg ccacgcccgc 81999
gtcgatgaac gcctcgatga ggaaatagac gaacgcgggc atggggcccg cgaggcccg 82059
cgccatctcc acgcccgcct cgtcggcgac ctccacttcg gtgccgagcg cgcccagcag 82119
tgagcggacg cgcaccagcg ccggctcgct caagcccgcg cttggaatcc agaagggtggc 82179
cgccttgccg acccgcgcg ggggtgttggg catcgcgcg aacgacgagt ggtgttcgag 82239
accctcgcac agcgtctgca ggcgcgcgc cgcgatgatg ctgacgaaa tcgtgtcctc 82299
acggagctgc ccgcgcagct gctggaggac aacaccaggt tgttgccggt tcacggccag 82359
cacgaccagt tccgcgtccg ccgcccgcg gttgtccgc gtcacctga cgccatactg 82419
ctcgccggagc acccggcgac gtcctcgtc gacctctcc acgaggagcg tgccggggcg 82479

ggccaggcct cccccggtga ccgctccaat caaggcctcg gccatcacc caccgccgat 82539
gaagcccagg gctggcctcc gcgagggcac ggtggcatcc atcgaatcgg ccgtcatcgc 82599
gaggtacgaa cgggattggg gcggatgagt cccttcacga tgcgaccgt ctcgacatcg 82659
gccagggcct gattcacctc ttcgagggcg taggtgcggc tgaccatgtg ctcccagccc 82719
ttgccaccgg cgaccctcgc accgtgccgg tcgaggatgc ggagcatccg gtggaagtgg 82779
ctgaagtcca cgccccacac acctcgaatc tcgaggtgct ttcgattgat cccttttagtg 82839
agggttaatt gcggccgcga attcgtaat 82868

<210> 2

<211> 481

<212> PRT

<213> Angiococcus disciformis

<400> 2

Val Thr Arg Pro Tyr Pro Ser Ser Leu Pro Arg Leu Gly Gly Lys Pro
1 5 10 15

Ser Ser Gln Met Ala Arg Ile Leu Leu Val Asp Asp Glu Lys Ile Ala
20 25 30

Arg Thr Leu Tyr Gly Asp Tyr Leu Thr Ala Val Gly His Ala Val Thr
35 40 45

Ala Val Gly Ser Leu Gln Asp Ala Arg Glu Ala Leu Ala Gly Asp Arg
50 55 60

Phe Asp Ala Val Val Thr Asp Leu Ile Leu Pro Gly Gly Asp Gly Met
65 70 75 80

Glu Val Leu Arg His Val Arg Glu Arg His Pro Gly Val Glu Val Val
85 90 95

Val Ile Thr Gly Leu Glu Lys Val Asp Pro Ala Val Arg Ala Ile Lys
100 105 110

Ser Gly Ala Ala Glu Tyr Leu Val Lys Pro Val Ala Pro Glu Ala Leu
115 120 125

Gln His Ala Val Arg Arg Ala Leu Thr Thr Arg Asp Leu Leu Gln Glu
130 135 140

Asn Ala Ser Leu Arg Arg His Val Ala Met Leu Glu Ala Gly Gln Arg
145 150 155 160

Ile Ala Thr Thr Leu Glu Arg Glu Lys Leu Ala Ser Ala Thr Thr Ser
165 170 175

Ala Leu Glu Ser Met Ala Cys Ala Ser Ala Val Val Leu Leu Glu Arg
180 185 190

Asp Gly Asp Lys Gly Leu Arg Leu Gln Gly Val Arg Gly Leu Pro Pro
195 200 205

Arg His Glu Pro Glu Leu Val Ala Thr Leu Val Glu Arg Leu Ser Asn
210 215 220

Ala Arg Ala Pro Arg Glu Leu Glu Gly Leu Asp Ala Ala Phe Pro Arg
225 230 235 240

Val Leu Ser Ile Pro Ala Val Glu Gly Thr Thr Val Leu Gly His Ala
245 250 255

Val Leu Phe Phe Asp Gly Ala Glu Pro Glu Trp Thr Gly Glu Thr Ala
260 265 270

Gly Tyr Leu Val Arg Asn Trp Ala Leu Ala Leu Arg Asn Leu Gly Arg
275 280 285

Phe Ala Ala Val Glu Asp Leu Ala Tyr Val Asp Asp Leu Thr Arg Leu
290 295 300

Phe Asn Thr Arg Tyr Leu His Leu Val Leu Asp Arg Glu Val Gln Asp
305 310 315 320

Ser Val Gln Thr Gln Arg Pro Phe Ser Leu Leu Phe Leu Asp Leu Asp
325 330 335

His Phe Lys Ser Ile Asn Asp Thr His Gly His Leu Val Gly Ser Arg
340 345 350

Leu Leu Val Glu Thr Ala Arg Val Val Lys Gly Cys Val Arg Asp His
355 360 365

Asp Val Val Ala Arg Tyr Gly Gly Asp Glu Tyr Val Val Leu Leu Arg
370 375 380

Asn Thr Asp Ser Gly Gly Ala Leu Lys Val Ala Glu Arg Ile Arg Arg
385 390 395 400

Thr Met Glu Thr His Leu Phe Leu Ala Arg Glu Gly Leu Ser Leu Lys
405 410 415

Leu Ser Thr Cys Ile Gly Val Ala Ser Phe Pro Glu His Ala Arg Asp
420 425 430

Lys Ala Thr Leu Leu Asp Leu Ala Asp Arg Ala Met Tyr Arg Gly Lys
435 440 445

Arg Gly Ser Arg Asn Val Val Tyr Met Ala Ala Leu Asp Leu Glu Pro
450 455 460

Leu Pro Ala Glu Arg Arg Gln Gly Gly Ser Gly Pro Gly Gly Ser Ser
465 470 475 480

Thr

<210> 3

<211> 346

<212> PRT

<213> *Angiococcus disciformis*

<400> 3

Val Gly Gly Val Glu Ala Arg Leu Asp Ala Val Ala Asp Ala Phe Glu
1 5 10 15

Ala Gly Asp Phe Glu Ala Ala Leu Ala Gly Ala Glu Gly Leu Leu Ala
20 25 30

Asp Ala Pro Glu Leu Pro Glu Ala Leu His Phe Arg Ala Ser Ala Leu
35 40 45

Val Glu Leu Gly Arg Leu Glu Glu Ala Gly Lys Ala Phe Gly Gln Ala
50 55 60

Leu Lys Val Ala Pro Glu Asp Leu Glu Ile Leu Leu Ser Ala Ala Asp
65 70 75 80

Cys Leu Val Cys Arg Ala Gly Glu Asp Arg Glu Ala Val Ala Glu Gly
85 90 95

Leu Ala Leu Cys Ala Arg Gly Arg Arg Leu Ala Gln Lys Ala Asp Asp

100

105

110

Val Glu Met Leu Tyr Glu Phe Leu Leu Leu Glu Gly Met Gly Leu Asn
115 120 125

Gln Met Gly Glu Cys Ala Thr Ala Leu Val Ser Leu Asp Ala Ala Leu
130 135 140

Gly His Met Pro Arg Ser Leu Asp Ala Gln Val Glu Arg Gly Ile Ala
145 150 155 160

Leu Phe Glu Leu Cys Arg Phe Asp Glu Ala Lys Ala Ala Phe Glu Lys
165 170 175

Val Leu Lys Asp Ala Pro Asp Asp Pro Trp Ala His His Tyr Leu Gly
180 185 190

Leu Ile Ala Glu Arg Arg Gly Asp Glu Lys Glu Ala Lys Arg Arg Phe
195 200 205

Asp Lys Ala Arg Ala Leu Val Pro Glu Glu Phe Pro Pro Pro Val Glu
210 215 220

Leu Ala Glu Ala Glu Phe Asp Arg Ala Val Glu Asp Ala Val Lys Ser
225 230 235 240

Leu Pro Arg His Ala Lys Gln Tyr Leu Asp Asn Val Thr Ile Ala Val
245 250 255

Glu Asp Leu Pro Ser Asp Glu Asp Leu Leu Gly Gln Asp Pro Pro Leu
260 265 270

Ser Pro Ser Ile Leu Gly Val Phe Arg Gly Thr Pro Val Gly Glu Arg
275 280 285

Ser Val Met Asn Ala Tyr Glu Leu Pro Ala Ser Ile Val Leu Tyr Gln
290 295 300

Arg Asn Leu Glu Arg Phe Ala Arg Thr Arg Glu Glu Leu Ile Glu Gln
305 310 315 320

Ile Gly Ile Thr Val Met His Glu Val Gly His Leu Met Gly Leu Asp
325 330 335

Glu Asp Asp Leu Trp Gln Arg Gly Leu Asp
340 345

<210> 4

<211> 214

<212> PRT

<213> Angiococcus disciformis

<400> 4

Met Gly Ile Ser Glu Arg Lys Glu Arg Gln Arg Ala Glu Leu Arg Glu
1 5 10 15

Gln Ile Leu Gln Val Ala Lys Asp Met Val Thr Arg Glu Gly Phe Gly
20 25 30

Ala Leu Ser Met Arg Lys Leu Ala Asp Ala Val Glu Tyr Ala Pro Ala
35 40 45

Thr Leu Tyr Leu His Phe Glu Asn Arg Asp Ala Ile Ala Arg Glu Leu
50 55 60

Cys Ile Arg Gly Phe Gln Asp Leu Leu Glu Ala Phe Glu Pro Ala Ala
65 70 75 80

Ser Val Glu Glu Pro Val Glu Arg Leu Tyr Arg Leu Gly Glu Ala Tyr
85 90 95

Val Lys Phe Gly Leu Glu Gln Pro Glu Thr Tyr Arg Leu Ile Phe Met
100 105 110

Glu Asp Pro Lys Leu Ser Thr Ala Leu Phe Arg Asp Ala Pro Glu Asp
115 120 125

Gly Ala Gly Pro Arg Ser Phe Gly Val Leu Val Lys Val Phe Glu Asp
130 135 140

Leu Lys Ala Ala Gly Arg Ile Ala Glu Asp Ala Glu Pro Ser Lys Leu
145 150 155 160

Ala Glu Val Leu Trp Ala Gly Val His Gly Ile Val Ala Leu Arg Leu
165 170 175

Thr Cys Thr Gly Phe Lys Gly Ser Pro Pro Glu Glu Leu Ala Arg Leu
180 185 190

Leu Val Ser Thr Met Val Asn Gly Leu Pro Gly Leu Lys Pro Val Lys
195 200 205

Ala Ala Arg Lys Thr Arg
210

<210> 5

<211> 378

<212> PRT

<213> *Angiococcus disciformis*

<400> 5

Met Thr Lys Trp Ser Gly Lys Trp Ile Gly Gly Arg Thr Tyr Thr Ala
1 5 10 15

Thr Asp Gly Ser Thr Arg Trp Ile Ile Arg Lys Thr Val Ala Gly Val
20 25 30

Ala Tyr Asn Val Thr Leu Asp Val Arg Ser Glu Ala Glu Ala Leu Ala
35 40 45

Glu Leu Ala Ala Phe Arg Arg Asn Pro Ala Ala Tyr Arg Thr Ala Ser
50 55 60

Gln Glu Arg Gln Asp Glu Ala Gln Gln Ala Gln Glu Asp Ala Ala Gln
65 70 75 80

Ala Val Phe Leu Asp Glu Asp Arg Ala Arg Arg Phe Leu Gln His Leu
85 90 95

Lys Ser Ser Gly Arg Ser Asp Glu Tyr Leu Lys Ser Thr Arg Ser Tyr
100 105 110

Leu Ala Ala Trp Ala Thr Ala Leu Ala Asp Lys Asp Leu Arg Ala Val
115 120 125

Thr Gly Pro Ala Leu Lys Lys Ile Leu Ala Thr Trp Asp Thr Gly Lys
130 135 140

Lys Trp Arg Ile Ile Val Phe Lys Ser Phe Thr Ser Phe Leu Gln Asp
145 150 155 160

Ser Gly Glu Leu Asp Pro Met Val Asn Pro Gly Arg Phe Leu Lys Val
165 170 175

Pro Pro Ala Arg Arg Asn His Glu Leu Lys Gly Tyr Ser Ile Glu His
180 185 190

Val Gln Gln Leu Tyr Ala Ala Leu Gly Ser Gln Ala Ile Arg Asp Val
 195 200 205

Leu Cys Leu Gln Ala Lys Thr Gly Met His Gly Thr Glu Val Asp Arg
 210 215 220

Leu Ala Ser Gly Asp Gly Lys Ile Ala Val Leu Lys Asp Gln Gly Glu
 225 230 235 240

Ile Ala Ala Thr Val Thr Phe Lys His Lys Thr Gly Lys Pro Phe Thr
 245 250 255

Leu Ser Leu Asp Val Gln Gly Leu Ala Ala Ala Gln Arg Leu Gln Ala
 260 265 270

Arg Gly Ser Ala Pro Asp Arg Gln Thr Ile Arg Glu Ala Val Glu Arg
 275 280 285

Val Cys Ala Arg Thr Ser Leu Glu Phe Val His Phe Gly Ala Ile Arg
 290 295 300

His Ser Phe Ala Thr Trp Leu Val Glu Gln Gly Ser Ile Tyr Asn Pro
 305 310 315 320

Lys Gly Gly Gly Leu Ser Leu Glu Ala Val Ala Gln Ala Leu Asn His
 325 330 335

Ser Ser Thr Arg Thr Thr Ala Leu His Tyr Val Ser Ala Thr Val Pro
 340 345 350

Pro Met Tyr Val Val Pro Ile Arg Leu Glu His Ser Asn Asp Pro Ser
 355 360 365

Glu Leu Arg Ile Ala Arg Ala Pro Ala Gly
 370 375

<210> 6

<211> 388

<212> PRT

<213> Angiococcus disciformis

<400> 6

Met Thr Asp Asp Ser Thr Pro Thr Asp Glu Thr Gln Ile Glu Ala Pro

1

5

88
10

15

Thr Pro Glu Asn Thr Ser Asp Leu Lys Gly Ala Arg Pro Arg Leu Arg
20 25 30

Phe Thr Ile Glu Thr Thr Thr Tyr Glu Thr Leu Arg Thr Leu Glu Arg
35 40 45

Arg Gly Asn Gly Glu Val Val Leu Leu Ala Glu Arg His Leu Pro His
50 55 60

Gly Leu Ala Gly Leu Val Thr Ile Lys Arg Leu Arg Asn Pro Val Thr
65 70 75 80

Phe Glu Arg Cys Gln Arg Leu Ile Glu Glu Val Gln Leu Ser Phe Arg
85 90 95

Leu His His Pro Ala Ile Ala Gln Val His His Leu Lys Ile His Ala
100 105 110

Asp Arg Pro His Val Ile Ala Glu Tyr Val Asp Gly Pro Thr Leu Asp
115 120 125

Thr Ile Ile Ser Leu Ala Thr Met Arg Glu Lys Pro Leu Ser Ala Pro
130 135 140

Phe Ala Leu Tyr Ile Ala Ala Glu Val Ala Asp Ala Leu His His Ala
145 150 155 160

His Thr Leu Arg Asp Ser Glu Asn Arg Pro Leu Gly Ile Ile His Arg
165 170 175

Asp Val Ala Pro Arg Asn Ile Arg Val Ala Arg Ser Gly Glu Val Lys
180 185 190

Val Thr Asp Phe Gly Ala Thr Tyr Ser Leu Met Val Gly Arg Glu Glu
195 200 205

Thr Pro Gly Leu Leu Leu Lys Gly Asp Val Ala Tyr Ala Ser Pro Glu
210 215 220

Tyr Leu Asn Arg Lys Pro Met Asp Gly Arg Ser Asp Ile Phe Ser Leu
225 230 235 240

Gly Leu Val Leu Met Glu Met Leu Thr Cys Lys His Leu Phe Asp Val
245 250 255

Glu Asp Glu Lys Ala Pro Asn Ala Thr Val Asp Val Lys Thr Glu Glu
260 265 270

Thr Pro Ser Val Pro Leu Thr Gln Met Ile Ala Leu Val Asn Arg Tyr
275 280 285

Arg Pro Glu Asp Val Glu Asn Ala Met Ala Gly Leu Pro Asp Ala Leu
290 295 300

Lys Ala Ile Val His Lys Ala Leu Gln Arg Lys Leu Ser Glu Arg Tyr
305 310 315 320

Ala Thr Ala Ala Glu Met Arg Asp Ala Leu Arg Ala Ala Leu Ala Ala
325 330 335

Glu Ser Gln Pro Phe Gly Arg Lys Glu Ala Ser Glu Glu Leu Ala Arg
340 345 350

Met Leu Ser Glu Ala Ser Val Leu Arg Asp Arg Val Glu Leu Asp Glu
355 360 365

Glu Gly Ile Phe Pro Glu Gly Leu Asp Ala Asp Glu Pro Thr Pro Ala
370 375 380

Pro Asn Glu Glu
385

<210> 7

<211> 640

<212> PRT

<213> *Angiococcus disciformis*

<400> 7

Val Glu Glu Lys Leu Gly Ala Gly Gly Phe Gly Ala Val Tyr Arg Ala
1 5 10 15

Arg Arg Gly Gly Arg Leu Tyr Ala Leu Lys Leu Ile Pro Leu Trp Gly
20 25 30

Leu Ala Glu Trp Ala Glu Arg Glu Val Ala Ile Leu Leu Arg Leu Lys
35 40 45

His Ser Asn Leu Val Arg Ile Arg Gly His Gly Gln Trp Pro Asp Glu
50 55 60

Ala Pro Gln Ser Phe Phe Ile Val Met Asp Tyr Val Glu Gly Arg Arg
65 70 75 80

Leu Asp Val Trp Ala Arg Glu Glu Asn Pro Ser Ala Arg Glu Val Val
85 90 95

Leu Lys Val Arg Gly Val Ala Arg Gly Leu Gly Ala Ala His Arg Ala
100 105 110

Lys Val Val His Arg Asp Leu Lys Glu Ser Asn Val Ile Glu Arg Thr
115 120 125

Ser Asp Gly Glu Ala Val Val Val Asp Phe Gly Ala Gly Gly Tyr Glu
130 135 140

Ser Ala Pro Ser Ile Thr Gly Gly Val Leu Pro Pro Gly Thr Pro Glu
145 150 155 160

Tyr Arg Ala Pro Glu Ala Trp Arg Phe Gln Gln Glu His Gly Asp Glu
165 170 175

Arg Gly His Ser Tyr Gln Pro Gly Pro Ser Asp Asp Leu Tyr Ser Leu
180 185 190

Gly Val Val Leu Tyr Trp Leu Leu Thr Gly Arg Gln Pro Phe Leu Pro
195 200 205

Asp Glu Ala Ala Gly Val Glu Ala Val Leu Asn Arg Ala Pro Lys Pro
210 215 220

Pro His Val Leu Asn Pro Arg Val Pro Glu Ala Leu Ser Ala Val Cys
225 230 235 240

Met Arg Leu Leu Ala Lys Glu Pro Glu Glu Arg His Pro Asp Ala Asp
245 250 255

Ala Leu Cys Ala Glu Leu Glu Ala Leu Leu Ala Gln Ala Asp Glu Ser
260 265 270

Trp Asp Val Lys Leu Cys Asp Ala Tyr Gly Pro Asp Thr Ala Thr Thr
275 280 285

Leu Ala Ala Val Pro His Ala Gly Glu Asp Glu Leu Val Gln Trp Leu
290 295 300

Lys Lys Arg Lys Ala Arg Pro Arg Arg Gly Pro Arg Pro Pro His Gly
305 310 315 320

Glu Gly Ala His Glu Pro Asp Ser Asn Ala Val Ala Ile Gln Ala Glu
325 330 335

Leu Pro Pro Pro Val Pro Ala Arg Pro Gly Pro Met Pro Ser Pro Pro
340 345 350

Pro Ala Gln Ala Gly His Phe Lys Ala Leu Arg Val Ala Thr Trp Met
355 360 365

Gly Leu Leu Val Val Leu Ser Ala Gly Ala Leu Val Ile Ala Arg Thr
370 375 380

Leu Ala Pro Ser Pro Ala Pro Ala Ala Arg His Ala Leu Glu Pro His
385 390 395 400

Pro Ser Asp Ala Val Asp Ser Pro Pro Ala Ser Pro Glu Phe Asp Leu
405 410 415

Ala Pro Trp Ala Pro Gly Gln Glu Val Ala Ala Pro Trp Met Ala Pro
420 425 430

Glu Ala Gln Glu Ala Ala Ala Ala Leu Asp Ser Ala Pro Thr Leu Ala
435 440 445

Ala Val Ala Leu Pro Ala Thr Ile Ser Glu Glu Lys Ala Ser Val Lys
450 455 460

Met Lys Lys Asn Thr Gly Val Phe Ser Glu Pro Glu Pro Gln Arg Phe
465 470 475 480

Arg Gly Gly Thr Val Gly Lys Ala Leu Gly Leu Gly Val Ala Gly Cys
485 490 495

Leu Ala Met Ala Cys Pro Gly Pro Gln Val Arg Pro Thr Pro Pro Pro
500 505 510

Glu Ala Cys Pro Pro Gly Ala Val Glu Ala Met Ser Arg Leu Arg Ile
515 520 525

Glu Tyr Asp Glu Asp Val Val Ala Thr Phe Phe Ala Thr Ala Glu Glu
530 535 540

Gly Val Ser Arg Asn Ile Thr Val Arg Ala Gly Pro Thr Thr Val His
545 550 555 560

Leu Gly Arg Pro Leu Gly Asn Leu Pro Ala Arg Thr Thr Leu Ser Gly

Arg Leu Ile Ile Gly Glu Asp Arg Val Tyr Gly Arg Leu Thr Glu Ala
580 585 590

Arg Thr Pro Lys Gly Asp Arg Phe Pro Val Cys Met Gln Ile Ile Glu
595 600 605

Asn Asp Ser Ile Gly Phe Lys Lys Glu Ile Asn Asp Gly Ser Asn Ser
610 615 620

Ala Ser Val Tyr Pro Tyr Phe Asp Val Lys Ala Val Arg Arg Phe Glu
625 630 635 640

<210> 8

<211> 306

<212> PRT

<213> *Angiococcus disciformis*

<400> 8

Met Ser Pro Ile Val Val Leu Leu Val Ile Ala Leu Phe Ser Gly Ala
1 5 10 15

Ala Ala Ala Gln Ser Arg Pro Ala Ala Thr Gly Leu Gly Val Arg Arg
20 25 30

Ile Glu Phe Ser Ala Glu Asn Phe Glu Ala Met Thr Thr Pro Glu Val
35 40 45

Gln Ile Ser Pro Gly Val Ser Thr Thr Phe Glu Phe Asn Ser Ala Leu
50 55 60

Leu Gln Glu Lys Val Ala Val Glu Gly Ala Asp Arg Phe Ser Leu Val
65 70 75 80

Asp Ile Gly Arg Ser Thr Leu Arg Leu Val Pro Ser Glu Gln Val Leu
85 90 95

Pro Gly Glu Arg Leu Arg Val Thr Val Arg Phe Leu Asp Gly Ala Ala
100 105 110

Pro Val Gly Met Ala Phe Ile Leu Val Ala His Pro Ala Gln Ala Glu
115 120 125

Arg Leu Val Glu Val Trp Arg Asn Gln Arg Thr Val Glu Ser Tyr Gln
 130 135 140

Gln Glu Ser Lys Glu Ala Arg Ala Glu Ala Gln Arg Cys His Glu Glu
 145 150 155 160

Asn Glu Arg Leu Arg Ala Glu Tyr Ala Gly Pro Gly Gly Leu Ala Gly
 165 170 175

Leu Leu Ala Asn Arg Val Ile Gly Lys Ser Gly Val Ser Val Lys Pro
 180 185 190

Leu Asp Phe Glu Lys Glu Val Arg Gln Arg Pro Gly Asp Ala Ile Arg
 195 200 205

Ile Leu Arg Ala Trp Ser Tyr Arg Ser Ala Asn Arg Val Ser Val Leu
 210 215 220

Met Asp Leu Asp Tyr Pro Glu Ala Ala Gln Arg Trp Ile Ala Gln Gly
 225 230 235 240

Ala Ser Leu Val Ser Lys Thr Gly Glu Thr Leu Ser Met Leu Pro Val
 245 250 255

Trp Gln Glu Ala Pro Val Ala Asn Ala Arg Ile Arg Arg Leu Val Val
 260 265 270

Glu Ala Glu Ala Glu Pro Asp Ala Ala Gln Asp Thr Phe Thr Leu Lys
 275 280 285

Leu Trp Glu Ser Asn Gly Leu Arg Thr Ile Thr Leu Ser Gly Val Thr
 290 295 300

Phe Pro
 305

<210> 9

<211> 180

<212> PRT

<213> Angiococcus disciformis

<400> 9

Met Thr Leu Tyr Glu Leu Phe Glu Ala Asp Asp Asp Arg Asp Val Pro
 5 10 15

Ala Asp Phe Pro Trp Leu Leu Ile Glu Ile Pro Glu Glu Gln Leu Arg
20 25 30

His Ala Ala Leu Leu Leu Asn Glu Gly Glu Trp Arg Pro Ser His Ile
35 40 45

Ser Gly Ile Trp Tyr Arg Val Asp Pro Glu Arg Pro Ala Gln Lys Gln
50 55 60

Gln Arg His Val His Val Ala Ala Lys Lys His Ile Lys Ile Pro Thr
65 70 75 80

Lys Gln Ala Ser Trp Asn Arg Asp Thr Thr Arg His Asp Arg Lys Thr
85 90 95

Phe Asn Ala Lys Leu Gly Ser Gln Gly Ser Tyr Gln Asp Val Ala Lys
100 105 110

Thr Ala Leu Gly Leu Pro Pro Glu Ala Val Leu Glu His Val Val Glu
115 120 125

Asn Pro Ala Glu Gln Gln Met Leu Leu Thr Glu Ser Thr Asp Pro Gln
130 135 140

Leu Lys Thr Glu Phe Tyr Arg Trp Leu Glu Ala Gln Pro Arg Arg Ser
145 150 155 160

Asn Phe Glu Ala Leu Val Glu Asp Ala Val Arg Leu Asn Leu Pro Leu
165 170 175

Asp Leu Glu Ser
180

<210> 10

<211> 390

<212> PRT

<213> *Angiococcus disciformis*

<400> 10

Val Ala Leu Gln Arg Arg His Pro Leu Leu Arg Ala Lys Ile Val Ala
1 5 10 15

Gly Thr Pro Pro Arg Phe Thr Ser Gln Gly Val Pro Pro Ile Ala Leu
20 25 30

Glu Val Val Glu Arg Lys Ala Glu Glu Ser Trp Gln Gln His Val Glu
 35 40 45

Ala Glu Leu Asp Arg Pro Phe Asn Ser Asp Pro Gly Pro Leu Val Arg
 50 55 60

Phe Ile Leu Val Arg Gly Glu His Arg Ser Glu Leu Leu Cys Ser Tyr
 65 70 75 80

Asp His Leu Ile Gly Asp Ala His Ser Gly Ile Phe Ala Leu Arg Asp
 85 90 95

Leu Leu Gln Val Met Ala Ser Gln Asp Gln His Leu Pro Glu Leu Ala
 100 105 110

Pro Arg Pro Ala Tyr Glu Glu Leu Ile Gly Pro Met Val Pro Gly Thr
 115 120 125

Arg Leu Leu Gly Ala Ala Val Arg Arg Gly Ser Ser Ala Leu Leu Arg
 130 135 140

Leu Ser Pro Thr Leu Asn Thr Leu Thr Glu Arg Leu Val Ala Pro Arg
 145 150 155 160

Gly Met Ala Arg Ser Ser Ala Glu Gly Gln Ala Leu Tyr Ser His Arg
 165 170 175

Ile Leu Glu Pro Glu Gln Leu Ala Arg Leu Leu Ala Arg Cys Arg Glu
 180 185 190

Gln Gly Ser Ser Val His Ala Ala Leu Gly Ala Ala Leu Leu Ile Ala
 195 200 205

Arg Ala Glu Ser Gln Gly Ala Lys Lys Arg Val Thr Leu Thr Leu Thr
 210 215 220

Ser Ala Leu Asp Ala Arg Glu Arg Phe Gly Val Gly Glu Asp Phe Gly
 225 230 235 240

Leu Phe Thr Thr Gly Lys Thr Asp Phe Phe Arg Ala Arg Gly Ser Thr
 245 250 255

Pro Phe Trp Glu Leu Ala His Arg Leu Arg Ser Pro Leu Gln Ala Ala
 260 265 270

Arg Lys Lys Arg Ser His Leu Arg Leu Phe Arg Leu Ile Leu Glu Ile

275

280

285

Ser Ala Leu Thr Val Asp Ile Ser Ser Gln Pro Trp Met Glu Arg Ala
290 295 300

Thr Arg Leu Gly Leu His Ser Met Leu Ala Leu Ser Asn Val Gly Arg
305 310 315 320

Val Glu Ile Ala Ala Arg Tyr Gly Asn Leu Thr Leu Glu Gly Leu Gly
325 330 335

Phe Thr Gly Thr Thr Ala Ala Gln Phe Asp Val Val Leu Thr Ala Ile
340 345 350

Thr Phe Ala Gln Arg Leu Glu Ala Ser Phe Leu Phe Asn Thr Ala His
355 360 365

Met Gln Arg Glu Gln Val Glu Gln Leu Ala Ser Arg Thr Trp Glu Leu
370 375 380

Leu Ala Glu Ala Thr Arg
385 390

<210> 11

<211> 1541

<212> PRT

<213> Angiococcus disciformis

<400> 11

Met Thr Leu Asp Ser Arg Tyr Ala Asn Arg Gly Gly Gly Leu Tyr Glu
1 5 10 15

Ile Asp Gly Leu Ala Ile Gly His Ser Pro Pro Pro Cys Ser Ile Pro
20 25 30

Gly Gly Val Gln Leu Thr Ala Thr Ala Thr Ala Glu His Ala Leu Lys
35 40 45

Ala Leu Arg Ala Lys Gly Leu Ala Thr Ser Asp Glu Ala Leu Pro Ala
50 55 60

Val Arg His Ala Ser Gly Glu Lys Ala Pro Leu Ser Ser Ser Gln Arg
65 70 75 80

Arg Met Trp Phe Leu Glu Gln Leu Glu Pro Gly Asn Ala Ala Gln His
85 90 95

Leu Leu Gln Ser His His Ile Gln Gly Pro Leu Gln Val Glu Gly Leu
100 105 110

Arg Arg Ala Leu Gly Leu Met Val Lys Arg His Glu Ala Leu Arg Leu
115 120 125

Val Val Ala Ile Gly Ala Ala Gly Pro Glu Gln Arg Val Arg Pro Glu
130 135 140

Trp Trp Pro Glu Leu Pro Leu Ser Asp Leu Ser Gly Val Ala Gln Ala
145 150 155 160

Asp Arg Ala Arg Ala Leu Ala Glu Leu Ala Arg Val Glu Ala Ala Ala
165 170 175

Pro Phe Asn Leu Gln Gln Gly Pro Leu Phe Arg Val Arg Leu Ala Arg
180 185 190

Leu Ala Glu Thr Glu His Val Leu Leu Val Thr Leu His His Leu Ile
195 200 205

Ser Asp Gly Ala Trp Ser Cys Glu Val Leu Ile Lys Glu Leu Ala Met
210 215 220

Leu Tyr Gly Gln His Val Glu Gly Ser Ser Pro Glu Leu Ala Pro Leu
225 230 235 240

Pro Val Gln Tyr Arg Asp Tyr Ala Gly Trp Glu Ala Ser Phe Ala Pro
245 250 255

Ala Gly Glu Ala Leu Glu Ala Trp Trp Arg Gln Arg Leu Ala Gly Val
260 265 270

Pro Thr Val Leu Glu Leu Pro Thr Glu Gly Ala Arg Pro Leu Arg Gln
275 280 285

Thr Tyr Arg Ala Gly Arg Val Ala Ile Thr Val Gln Pro Arg Leu Arg
290 295 300

Gln Ala Leu Glu Asp Leu Ala His Lys Glu Gly Val Ser Leu Phe Ala
305 310 315 320

Leu Leu Leu Thr Ala Phe Thr Thr Leu Leu His Arg Tyr Ser Arg Gln
325 330 335

Glu Glu Leu Val Val Gly Trp Pro Ala Pro Gln Arg Pro Arg Pro Glu
340 345 350

Leu His Gly Leu Ile Gly Tyr Phe Gly Thr Pro Val Ala Leu Arg Ser
355 360 365

Arg Leu Glu Pro His Thr Arg Val Arg Glu Ala Leu Arg Gln Leu Asp
370 375 380

Gln Glu Val Arg Glu Ala Asn Ala His Ala Ala Leu Pro Phe Glu Arg
385 390 395 400

Leu Val Ser Leu Leu Asp Ile Ala Arg Ser Pro Ser Arg His Pro Leu
405 410 415

Phe Gln Val Leu Phe Asp Leu Leu Pro Glu Gln Pro Gln Pro Thr Ala
420 425 430

Gly Gly Cys Ala Phe Arg Pro Trp Glu Ser Phe Thr Gly Leu Val Ala
435 440 445

Tyr Asp Leu Thr Leu Leu Leu Glu Pro Arg Gly Glu Gly Leu Glu Gly
450 455 460

Ala Leu Asp Tyr Ser Ala Asp Leu Phe Ser Glu Ala Arg Met Gln Arg
465 470 475 480

Ala Ala Thr Gln Tyr Leu His Leu Leu Glu Gln Leu Val Glu Arg Pro
485 490 495

His Glu Arg Leu Ser Arg Leu Ala Leu Leu Thr Pro Gly Glu His Glu
500 505 510

Ala Leu Leu Ala Gly His Gly Leu Gly Gly Pro Val Glu Gly Ala Gln
515 520 525

Val Leu Ala His His Arg Phe Glu His Gln Val Arg Leu Thr Pro His
530 535 540

His Pro Ala Leu Cys Phe Gly Pro Gln Val Leu Ser Tyr Glu Gln Leu
545 550 555 560

Asn Arg Arg Ala Asn Pro Leu Ala His Arg Leu Arg Arg Leu Gly Ala
565 570 575

Gly Pro Asp Thr Leu Val Gly Leu Cys Val Glu Arg Ser Leu Glu Leu
580 585 590

Pro Val Ala Leu Leu Ala Ile Trp Lys Ala Gly Ala Gly Phe Leu Pro
595 600 605

Leu Asp Val Asn Gln Pro Arg Glu Arg Leu Ala Phe Leu Leu Gly Asp
610 615 620

Ala Ser Cys Arg Ile Leu Leu Thr Gln Glu His Leu Leu Gln Arg Leu
625 630 635 640

Pro Pro Thr Asn Ala Ala Leu Leu Cys Leu Glu Arg Glu Ala Glu Ala
645 650 655

Leu Glu Arg Glu Pro Gln Glu Asp Ala Pro His Glu Ala Gly Leu Asp
660 665 670

Asn Leu Ala Tyr Val Ile His Thr Ser Gly Ser Thr Gly Thr Pro Lys
675 680 685

Gly Ile Ala Met Val His Arg Cys Leu Ala Asn Leu Val Ala Trp Gln
690 695 700

Leu Thr His Glu Arg Leu Gly Gly Pro Ser Arg Thr Leu Gln Phe Ala
705 710 715 720

Ser Leu Asn Phe Asp Ile Cys Tyr Gln Glu Leu Phe Thr Thr Trp Ala
725 730 735

Ala Gly Gly Thr Val Val Met Val Thr Glu Glu Val Arg Arg Asp Pro
740 745 750

Ala Arg Leu Leu Glu Val Leu Glu Gln Glu Gln Val Ser Arg Leu Tyr
755 760 765

Leu Pro Phe Ile Ala Leu Gln Gln Leu Ala Arg Val Ala Asp Glu Arg
770 775 780

Gly Ala Ala Pro Arg His Leu Arg Gln Leu Ile Thr Ala Gly Glu Gln
785 790 795 800

Leu Gln Ala Thr Pro Glu Leu Gln Arg Leu Leu Ser Arg Met Pro Glu
805 810 815

Cys Thr Leu His Asn Gln Tyr Gly Pro Ser Glu Cys His Val Val Thr
820 825 830

Ser His Asp Leu Thr Arg Glu Pro Ser Arg Trp Pro Arg Leu Pro Pro

835

840

845

Val Gly Arg Pro Leu Ala His Leu Arg Val Leu Leu Leu Asp Gly Glu
850 855 860

Gln Gln Leu Val Pro Pro Gly Val Ala Gly Glu Val Phe Leu Gly Gly
865 870 875 880

Pro Ala Leu Ala Arg Gly Tyr Leu Gly Arg Pro Glu Gln Thr Ala Asp
885 890 895

Arg Phe Val Pro Asp Pro Phe Ser Arg Glu Pro Gly Ala Arg Leu Tyr
900 905 910

Arg Thr Gly Asp Leu Ala Arg Leu Arg Glu Asp Gly Ala Leu Glu Phe
915 920 925

Leu Gln Arg Met Asp Ala Gln Val Lys Ile Arg Gly Tyr Arg Ile Glu
930 935 940

Pro Gly Glu Ile Glu Val Val Leu Cys Glu His Pro Ala Val His Gln
945 950 955 960

Ala His Val Arg Pro Tyr Val Asp Ser Ala Gly Glu Arg Arg Leu Val
965 970 975

Ala Tyr Val Ala Ala Arg Leu Glu Asp Thr Asp Gly Ala Glu Thr Glu
980 985 990

His Val Glu Arg Trp Arg Ala Val Trp Asp Glu Thr Tyr Gly Gly Pro
995 1000 1005

Ser Gly Thr Ala Phe Asp Leu Ala Gly Trp Asn Asp Ser Val Arg
1010 1015 1020

Gly Glu Pro Leu Pro Pro Glu Gln Met Arg Glu Trp Val Glu Thr
1025 1030 1035

Thr Val Glu Arg Leu Met Glu Leu Val Pro Arg Arg Val Leu Glu
1040 1045 1050

Leu Gly Cys Gly Ser Gly Leu Leu Leu Arg Arg Leu Ala Pro Arg
1055 1060 1065

Cys Glu Ser Tyr Trp Gly Thr Glu Leu Ser Pro Val Ala Val Glu
1070 1075 1080

Arg Leu Arg Glu Gln Leu Gln Thr Gly Gly Ser Pro Leu Ala Gln
1085 1090 1095

Arg Val Arg Leu Met Ala Gln Pro Ala Asp Asp Phe Ser Gly Leu
1100 1105 1110

Pro Glu Ala Gly Phe Asp Thr Val Ile Leu Asn Ser Val Thr Gln
1115 1120 1125

Leu Phe Pro Ser Val Asp Tyr Leu Leu Arg Val Val Glu Gly Ala
1130 1135 1140

Leu Arg Val Leu Gln Pro Gly Gly Thr Leu Phe Ile Gly Asp Val
1145 1150 1155

Gln Asn Leu Arg Leu Phe Glu Leu Phe His Ala Ser Val Ala Leu
1160 1165 1170

Glu Gln Ala Ser Ala Asp Leu Glu Ala Pro Ala Leu Leu Ala Arg
1175 1180 1185

Thr Arg Gln Arg Met Leu Leu Asp Glu Arg Leu Tyr Val Asp Pro
1190 1195 1200

Asp Phe Phe Ala Ala Leu Ala Thr His Phe Pro Gln Leu Gly Ala
1205 1210 1215

Val Arg Leu His Leu Lys Arg Gly Ser Gly Arg Asn Glu Met Asn
1220 1225 1230

Arg Phe Arg Tyr Asp Val Glu Leu Gln Leu Ala Pro Val Ala Lys
1235 1240 1245

Ala Gly Pro Ala Gln Glu Leu Pro Glu Leu Asp Trp Arg His Glu
1250 1255 1260

Gly Leu Gly Leu Glu Arg Leu Glu Arg Met Leu Ala Glu Arg Pro
1265 1270 1275

Ala Gly Leu Val Leu Arg Asn Val Ala Asn Ala Arg Thr Ala Asp
1280 1285 1290

Glu Ala Ala Arg Leu Ala Leu Leu Arg Thr Gly Asn Ser Val Gly
1295 1300 1305

Arg Leu Arg Ala Leu Pro Ala Val Thr Ser Trp Asn Pro Glu Gln
1310 1315 1320

Leu Trp Arg Leu Ala Glu Ala Ala Asp Tyr Thr Cys His Val Thr
1325 1330 1335

Trp Ser Ala Gln Asp Glu Glu Gly Arg Phe Asp Ala Leu Leu Met
1340 1345 1350

Ala Arg Ala Ala Gly Ser Ser Arg Pro Ala Ala Trp Leu Thr Pro
1355 1360 1365

Pro Pro Pro Pro Pro Arg Pro Trp Lys Ser Tyr Ala Asn Gln Pro
1370 1375 1380

Leu Ala Ala Ser Arg Arg Arg Thr Leu Val Gly Val Leu Arg Ser
1385 1390 1395

His Leu Glu His Lys Leu Pro Glu Tyr Met Val Pro Ser Ser Phe
1400 1405 1410

Val Leu Leu Asp Ala Leu Pro Leu Lys Pro Thr Gly Lys Leu Glu
1415 1420 1425

Val Ala Ala Leu Pro Pro Pro Glu Pro Ala Gln Ala Glu Gln Thr
1430 1435 1440

Pro Gly His Leu Ala Pro Arg Thr Pro Thr Glu Arg Arg Leu Ala
1445 1450 1455

Glu Leu Trp Arg Arg Val Leu Gly Val His Arg Val Gly Val Glu
1460 1465 1470

Asp Asn Phe Phe Gln Leu Gly Gly His Ser Leu Leu Ala Thr Arg
1475 1480 1485

Leu Leu Ser Leu Ile Arg Gly Glu Leu Gly Leu Glu Leu Pro Leu
1490 1495 1500

Arg Val Leu Phe Glu Gln Pro Thr Leu Ala Ala Met Ala Gly Cys
1505 1510 1515

Leu Glu Ala Gln Ser Trp Ser Thr Gln Ala Pro His Pro Pro Ser
1520 1525 1530

Ala Ser Ile Glu Glu Gly Glu Leu
1535 1540

<211> 3515

<212> PRT

<213> *Angiococcus disciformis*

<400> 12

Met Thr Pro Ser Gly Asp Glu Ala Leu Gln Ser Ser Ile Ala Leu Val
1 5 10 15

Gly Met Ala Gly Arg Phe Pro Gly Ala Pro Asp Val Glu Ser Phe Trp
20 25 30

Arg Asn Leu Val Ala Gly Val Glu Ser Ile Ser Phe Phe Ser Glu Glu
35 40 45

Glu Leu Arg Gln Ala Gly Val Ser Glu Gln Ile Arg Arg Arg Pro Glu
50 55 60

Tyr Val Pro Ala Lys Gly Val Leu Glu Asp Leu Glu Leu Phe Asp Ala
65 70 75 80

Gly Phe Phe Gly Tyr Ser Pro Arg Glu Ala Ser His Leu Asp Pro Gln
85 90 95

Gln Arg Leu Leu Leu Glu Cys Ser Trp Glu Ala Leu Glu Asp Ala Gly
100 105 110

Leu Arg Pro Asp Gln Leu Pro Gly Trp Val Gly Val Tyr Val Gly Ala
115 120 125

Gly Asp Thr Ser Tyr Arg Phe Gln Leu Leu Arg Gly His Gly Asp Pro
130 135 140

Leu Ser Gly Ser Lys Asp Val Ala Gly Phe Phe Gly Asn Tyr Pro Asp
145 150 155 160

Phe Leu Ala Thr Arg Val Ala Tyr Lys Leu Asn Leu Arg Gly Pro Ala
165 170 175

Leu Gly Ile His Thr Ala Cys Ser Thr Ser Leu Val Ser Ile Asn Met
180 185 190

Ala Cys Ser Ala Leu Arg Gly Phe Glu Cys Asp Met Ala Leu Ala Gly
195 200 205

Gly Val Ser Leu Arg Leu Pro Ala Arg Ser Gly Tyr Leu Tyr Glu Glu

210

215

220

Gly Gly Val Ala Ser Lys Asp Gly His Cys Arg Pro Phe Asp Ala Arg
 225 230 235 240

Ala Thr Gly Thr Val Thr Gly Asp Gly Val Gly Val Val Val Leu Lys
 245 250 255

Arg Leu Glu Asp Ala Leu Lys Ala Arg Asp Pro Ile His Ala Val Ile
 260 265 270

Arg Gly Trp Ala Leu Asn Asn Asp Gly Ala Ser Arg Ala Gly Phe Thr
 275 280 285

Ala Pro Ser Val Glu Gly Gln Ser Glu Val Ile Ala Leu Ala His Ala
 290 295 300

Ala Ala Gly Ile Ser Ala Arg Asp Ile Thr Tyr Val Glu Ala His Gly
 305 310 315 320

Thr Gly Thr Pro Leu Gly Asp Pro Ile Glu Val Ala Ala Leu Thr Arg
 325 330 335

Ala Phe Arg Ala His Thr Ala Asp Thr Ala Phe Cys Thr Leu Gly Ala
 340 345 350

Val Lys Ser Asn Ile Gly His Leu Asp Ala Ala Ala Gly Val Ala Gly
 355 360 365

Val Ile Lys Thr Val Gln Ala Leu Arg His Arg Leu Ile Pro Pro Thr
 370 375 380

Leu His Phe Glu Arg Pro Asn Pro Ala Leu His Leu Glu Gln Ser Pro
 385 390 395 400

Phe Phe Val Asn Thr Gln Pro Leu Pro Trp Glu Ser Pro Arg Gly Pro
 405 410 415

Arg Leu Ala Gly Val Ser Ser Phe Gly Ile Gly Gly Thr Asn Ala His
 420 425 430

Thr Leu Phe Glu Glu Ala Pro Pro Pro Pro Ala Ser Gly Pro Thr Arg
 435 440 445

Asp Gln Val Leu Leu Leu Ser Ala Arg Ser Thr Ser Ala Leu Glu
 450 455 460

His Ile Ala Gly Arg Leu Ala Ala His Leu Arg Arg His Pro Asp Leu
 465 470 475 480

Glu Leu Ala Asp Val Ala Phe Thr Leu Gln Val Gly Arg Ala Arg Phe
 485 490 495

Pro Tyr Arg Arg Ala Leu Thr Cys Arg Thr Leu Ala Glu Ala Met Glu
 500 505 510

Arg Leu Glu Ala Pro Glu Pro Arg Pro Pro Glu Pro Leu Ala His Glu
 515 520 525

Gly Glu Arg Pro Pro Leu Val Met Leu Phe Pro Gly Gln Gly Thr Pro
 530 535 540

Leu Val Gly Thr Ala Arg Ala Leu His Glu Ser Glu Pro Thr Phe Arg
 545 550 555 560

Gln Ala Val Glu Gln Cys Ala Arg Leu Leu Arg Gln Thr Leu Gly Leu
 565 570 575

Asp Val Arg Glu Val Leu Phe Pro Ser Ala Glu Gln Glu Glu Gln Ala
 580 585 590

Arg Arg Leu Ala Ala Gln Thr Arg Val Ala Gln Pro Ala Leu Phe Thr
 595 600 605

Leu Glu Tyr Ala Leu Ala Gln Thr Trp Leu Gly Trp Gly Leu Gln Pro
 610 615 620

Gln Ala Leu Ala Gly His Ser Leu Gly Glu Leu Val Ala Ala Cys Leu
 625 630 635 640

Ala Gly Val Phe Ser Leu Glu Asp Ala Leu Gln Leu Val Ala Ala Arg
 645 650 655

Gly Gln Leu Met Gln Gly Cys Pro Pro Gly Ala Met Leu Ala Val Pro
 660 665 670

Leu Pro Glu Ala Glu Leu Ala Ala Leu Leu Gly Ser Glu Leu Cys Ile
 675 680 685

Ala Ala Val Asn Gly Pro Arg Ala Cys Val Ala Ser Gly Pro Leu Pro
 690 695 700

Ala Val Glu Ala Leu Thr Ala Ala Leu Glu Ser Arg Gly Val Ser Ser
 705 710 715 720

Arg Arg Leu Glu Thr Ser His Ala Phe His Ser Ala Ser Met Glu Ala
 725 730 735

Cys Gln Gly Pro Leu Thr Thr Leu Leu Arg Arg Met Arg Leu Gln Ala
 740 745 750

Pro Arg Leu Pro Cys Val Ser Gly Leu Thr Gly Arg Trp Leu Thr Gly
 755 760 765

Glu Glu Ala Thr Glu Pro Thr Tyr Trp Ala Arg Gln Leu Arg Glu Pro
 770 775 780

Val Arg Phe Ser Glu Ala Leu Glu Thr Leu Trp Ser Leu Lys Glu Pro
 785 790 795 800

Val Leu Leu Glu Val Gly Pro Gly Thr Thr Leu Thr Ala Leu Ala Arg
 805 810 815

Arg His Pro Thr Arg Pro Ala Arg Thr Gln Glu Val Ala Ser Leu Pro
 820 825 830

Val Gln Pro Asp Thr Ala Val Pro Cys Ile Glu Glu Ala Val Gly Glu
 835 840 845

Leu Trp Gln Ala Gly Leu Glu Leu Asp Trp Ser Ala Leu His Ala Ala
 850 855 860

Pro Arg His Arg Ala His Leu Pro Pro Tyr Pro Phe Glu Arg Gln Arg
 865 870 875 880

Tyr Trp Ile Glu Pro Glu Ala Ala Pro Gln Pro Arg Ala Gln Gln Pro
 885 890 895

Thr Pro Ala Ser Leu Val Pro Pro Glu Gln Pro Ser Arg Glu Ala Leu
 900 905 910

Glu Asp Trp Phe Tyr Val Pro Thr Trp Glu Gln Ala Pro Ala Thr Ser
 915 920 925

Gly Gly Gly Gln Pro Leu Ala Gly Pro Val Leu Ala Phe Met Asp Ser
 930 935 940

Ser Gly Leu Ala Glu Gln Val Leu Ala Ala Leu Trp Pro Ala Asp Ser
 945 950 955 960

Gly Ala Leu Leu Thr Arg Val Glu Pro Ala Gly His Tyr Glu Gln Leu
 965 970 975

Ser Glu His Ala Phe Arg Leu Arg Pro Glu Ser Glu Glu Asp Trp Asp
 980 985 990

Ala Leu Phe Gln Ala Leu Gln Ser Gln Gly Arg Leu Pro Arg Arg Ile
 995 1000 1005

Leu His Ala Trp Ala Leu Thr Ala Glu Pro Gly Pro Cys Thr Pro
 1010 1015 1020

Asp Gly Glu Ala Val Leu Glu Gln Gly Phe Phe Ser Leu Leu Arg
 1025 1030 1035

Leu Ala Arg Ala Leu Gly Arg His Ala Pro Glu Arg Pro Val Gln
 1040 1045 1050

Leu Glu Val Leu Ser Ser Phe Val His Ala Val Gly Pro Arg Glu
 1055 1060 1065

Pro Leu Glu Pro Leu Lys Ala Thr Leu Leu Gly Ala Cys Ala Val
 1070 1075 1080

Leu Pro Leu Glu Tyr Pro His Val Gln Cys Arg Thr Ile Asp Val
 1085 1090 1095

Arg Pro Gly Ser Glu Pro Arg Glu Val Leu Val Arg Ser Leu Ala
 1100 1105 1110

Ala Glu Leu Ala Ala Pro Met Gly Glu Ser Pro Val Ala Trp Arg
 1115 1120 1125

Asp Gly Gln Arg Tyr Val Arg Arg Ala Thr Arg Gln Arg Leu Glu
 1130 1135 1140

Ala Ser Arg Pro Leu Arg Ser Leu Arg Glu Arg Gly Val Tyr Leu
 1145 1150 1155

Val Ala Gly Gly Leu Gly Gly Ile Gly Leu Val Leu Ala Arg Ala
 1160 1165 1170

Leu Ala Gln Arg Ala Arg Ala Arg Leu Ala Leu Leu Thr His Ser
 1175 1180 1185

Pro Phe Pro Pro Arg Glu Gln Trp Glu Gln Trp Leu Glu Glu Ala
 1190 1195 1200

Pro Ala His Pro Glu Pro Ala Trp Arg Ser Glu Ala Asp Pro Ser

1205		1210		1215
Glu Arg Arg Arg Thr Gln His	Arg Ile Arg Cys Leu	Leu Glu Leu		
1220	1225	1230		
Glu Gln Leu Gly Ala Glu Val	Gln Val Tyr Thr Ala	Asp Val Ala		
1235	1240	1245		
Glu Glu Ala Ala Val Arg Ser	Val Val Glu Gln Val	His Ala Arg		
1250	1255	1260		
Trp Gly Lys Ile His Gly Val	Leu His Ala Ala Ala	Thr Phe Asp		
1265	1270	1275		
Asp Gly Val Ile Gln Leu Arg	Thr His Glu Gln Ser	Ser Arg Ala		
1280	1285	1290		
Leu Arg Thr Lys Val Arg Gly	Ser Met Val Leu His	Glu Val Leu		
1295	1300	1305		
Ala Ser Glu Gly Leu Asp Trp	Phe Ala Leu Cys Ser	Ser Leu Ala		
1310	1315	1320		
Ser Ala Leu Gly Ser Phe Gly	Gln Ala Asp Tyr Cys	Ala Ala Asn		
1325	1330	1335		
Ala Phe Gln Asp Ala Tyr Ala	His His Leu Arg Arg	Gln Gly Phe		
1340	1345	1350		
Thr Gly Ala Leu Ala Leu Asp	Trp Gly Thr Trp Arg	Asp Thr Gly		
1355	1360	1365		
Ala Ala Met Arg Leu Val Ala	Arg Thr Arg Arg Gly	Gly His Glu		
1370	1375	1380		
Lys Pro Pro Thr Pro Leu Thr	His Pro Leu Phe Asp	Cys Glu Gln		
1385	1390	1395		
Arg Glu Pro Gly Gly Thr His	Trp Leu Gly Leu Thr	Leu Arg Gly		
1400	1405	1410		
Gly Glu Asp Trp Val Val Asp	Glu His Arg Leu Gln	Gly Val Pro		
1415	1420	1425		
Thr Leu Pro Gly Val Ala Tyr	Leu Glu Leu Ala Arg	Ala Ala Cys		
1430	1435	1440		

Ala Gln	Ala Leu Gly	Ala Glu	Ala Val Glu	Leu Ala	Glu Leu Leu
1445		1450		1455	
Leu Leu	Glu Pro Leu Thr	Val	Pro Arg Gly	Glu Ser	Arg Gln Val
1460		1465		1470	
Arg Val	Val Leu Gln Pro	Glu	Gly Gln Ala His	Ala	Leu Arg Val
1475		1480		1485	
Glu Ser	Arg Ser Glu Glu	Ala	Arg Gly Trp Asn	Glu	His Ala Arg
1490		1495		1500	
Gly Arg	Val Arg Ala Val	Pro	Arg Leu Ala Glu	Arg	Ile Gln Pro
1505		1510		1515	
Glu Leu	Leu Arg Ala Ala	Cys	Glu His Glu Gln	Pro	Val Pro Gly
1520		1525		1530	
Glu Pro	Gln Glu Gln Gly	Pro	Val His Ala Gly	Ala	Arg Trp His
1535		1540		1545	
Gly Leu	Phe Gln Trp Val	Arg	Arg Gly Pro Arg	Gln	Ala Leu Ala
1550		1555		1560	
Gln Leu	Ala Leu Pro Glu	Pro	Phe His Gly Asp	Leu	Glu Arg Phe
1565		1570		1575	
Glu Leu	His Pro Ala Leu	Met	Asp Met Ala Thr	Ser	Phe Ala Ile
1580		1585		1590	
Pro Gly	Gly Val Pro Trp	Leu	Ala Phe Gly Tyr	Glu	Arg Val Leu
1595		1600		1605	
Ile His	Gly Pro Leu Pro	Pro	Gln Val Leu Ser	His	Val Ser Leu
1610		1615		1620	
Pro Glu	Glu Ser Gln Ala	Gly	Ala Gln Gln Leu	Arg	Leu Gln Val
1625		1630		1635	
Arg Leu	Leu Asp Leu Glu	Gly	Trp Glu Arg Val	Arg	Ile Asp Gly
1640		1645		1650	
Tyr Leu	Leu Arg Pro Leu	Lys	Pro Ser Asp Ala	Ser	Val Glu Pro
1655		1660		1665	
Ala Ala	Pro Asn Val Glu	Val	Ala Val Gly Thr	Pro	Gly Leu Leu
1670		1675		1680	

Glu Ser Leu Gly Leu Arg Arg Cys Thr Arg Pro Ala Pro Gly Pro
 1685 1690 1695

Arg Gln Val Glu Ile Glu Val Glu Ala Ala Gly Leu Asn Phe Leu
 1700 1705 1710

Asp Val Leu Gly Ala Leu Gly Met Met Pro Ala Leu Glu Ala Glu
 1715 1720 1725

Glu Ser Val Leu Gly Arg Glu Cys Ser Gly Arg Ile Ala Ala Val
 1730 1735 1740

Gly Glu Gly Val Ser Gly Leu Arg Val Gly Asp Glu Val Leu Ala
 1745 1750 1755

Val Ala Pro Gly Cys Phe Arg Ser Tyr Val Leu Val Asp Glu Ser
 1760 1765 1770

Gln Val Val Arg Arg Pro Ala Ser Leu Gly Leu Ala Glu Gly Ala
 1775 1780 1785

Ala Gln Met Val Pro Phe Ala Thr Ala Tyr Phe Ala Leu His Thr
 1790 1795 1800

Val Gly Arg Leu Arg Arg Gly Glu Arg Ile Leu Ile His Ala Ala
 1805 1810 1815

Ala Gly Gly Leu Gly Leu Ala Ala Val Gln Leu Ala Ser Arg Thr
 1820 1825 1830

Gly Ala Glu Ile Leu Ala Thr Ala Gly Ser Glu Gln Lys Arg Glu
 1835 1840 1845

Tyr Leu Arg Ser Leu Gly Ile Ala His Val Leu Asp Ser Arg Ser
 1850 1855 1860

Thr Ser Phe Val Ser Glu Val Arg Glu Arg Thr Gly Gly Arg Gly
 1865 1870 1875

Val Asp Val Val Leu Asn Ser Leu Ala Gly Glu Leu Leu Leu Ala
 1880 1885 1890

Gly Leu Ser Val Leu Ala Pro His Gly Arg Phe Leu Glu Leu Gly
 1895 1900 1905

Asn Arg Asp Leu Tyr Ala Asp Gln Gln Val Gly Leu Arg Thr Leu
 1910 1915 1920

Ala Arg Gly Gln Thr Phe Ala Ala Ile Asp Phe Gly Pro His His
 1925 1930 1935
 Pro Asp Phe Arg Ala Val Leu Glu Glu Val Ala Thr Gln Leu Thr
 1940 1945 1950
 Gln Gly Gln Leu Glu Pro Leu Pro Thr Arg Leu Phe Pro Ala Arg
 1955 1960 1965
 Gln Val Ala Glu Ala Phe Ser Phe Met Ala Arg Ala Leu His Ile
 1970 1975 1980
 Gly Arg Val Ala Val Ser Met Gln Gly Ala Thr Ala Leu Pro Ala
 1985 1990 1995
 Ser Met Thr Arg Gly Ser Arg Pro Ala Pro Val Ala Val Pro Pro
 2000 2005 2010
 Trp Glu Asp Pro Arg Leu Ala Gly Gly Ile Ser Ser Glu Glu Gly
 2015 2020 2025
 Ala Glu Ala Phe Leu Arg Ala Leu Glu Gln Gly Ala Pro Gln Leu
 2030 2035 2040
 Ile Ile Ser Pro Gln Asp Phe Ser Ser Leu Leu Arg Gly Leu Gly
 2045 2050 2055
 Gly Ser Gln Gly Val Arg Glu Lys Glu Arg Leu Val Thr Gly Arg
 2060 2065 2070
 Ala Ala Ala Ala Glu Pro Gln Ala Leu Pro Pro Ser Ser Leu Glu
 2075 2080 2085
 Gln Leu Ile Glu Gln Val Trp Arg Lys His Leu Gly Val Glu Arg
 2090 2095 2100
 Val Gln Pro Thr Asp Ser Phe Phe Gln Leu Gly Gly Asp Ser Leu
 2105 2110 2115
 Leu Gly Ile Gln Val Ala Ala Asp Leu Arg Arg His Leu Gly Val
 2120 2125 2130
 Glu Leu Pro Thr Ala Thr Leu Phe Ser His Pro Thr Leu Ala Ala
 2135 2140 2145
 Leu Ala Ala Ala Leu Arg Ala Arg Gln Gly Glu Ala Ala Ala Pro

2150

2155

2160

Thr Ala Pro Ala Pro Ala Leu Val Pro Asp Pro Ala Ala Arg Phe
 2165 2170 2175

Glu Pro Phe Pro Leu Thr Asp Val Gln Glu Ala Tyr Trp Val Gly
 2180 2185 2190

Arg Arg Ser Ala Phe Glu Leu Gly Gly Val Ala Ala His Gly Tyr
 2195 2200 2205

Phe Glu Ile Glu Ser Pro Gly Leu Glu Val Glu Arg Phe Ile Gln
 2210 2215 2220

Cys Trp Arg Gln Leu Leu Gln Arg His Asp Met Leu Arg Met Val
 2225 2230 2235

Val Leu Pro Asp Gly Arg Gln Gln Val Leu Glu Gln Val Pro Glu
 2240 2245 2250

Tyr Thr Pro Glu Val Val Glu Leu Arg Gly Leu Ser Pro Gln Glu
 2255 2260 2265

Ala Glu Ser Arg Arg Leu Gln Leu Arg Glu Arg Met Ala His Gln
 2270 2275 2280

Val Leu Arg Ser Asp Arg Trp Pro Leu Phe Glu Leu Val Leu Cys
 2285 2290 2295

Arg Tyr Glu Gly Gly Val Arg Ile His Met Ser Met Asp Ala Leu
 2300 2305 2310

Met Leu Asp Ala Trp Ser Ser Ala Val Leu Arg Gln Asp Phe Ala
 2315 2320 2325

Gln Leu Tyr His Glu Pro Gly Arg Pro Leu Glu Pro Leu Ala Ile
 2330 2335 2340

Thr Phe Arg Asp Tyr Val Leu Ala Glu Arg Arg Leu Arg Glu Gly
 2345 2350 2355

Glu Ala His Glu Arg Ala Arg Ala Tyr Trp Trp Ala Arg Leu Asp
 2360 2365 2370

Thr Leu Pro Pro Pro Pro Glu Leu Pro Leu Val Lys Glu Pro Ser
 2375 2380 2385

Gln Leu Glu His Ala Arg Phe Thr His Arg Glu Ala Arg Leu Glu
2390 2395 2400

Pro His Arg Trp Ala Arg Leu Gln Glu Arg Ala Arg Ala His Gly
2405 2410 2415

Leu Thr Pro Ser Ala Ala Cys Met Ala Ala Phe Ala Glu Val Leu
2420 2425 2430

Ala Arg Trp Ser Arg His Pro Arg Phe Thr Leu Asn Leu Thr Leu
2435 2440 2445

Phe Gln Arg Leu Pro Leu His Pro Gln Val Asp Glu Leu Val Gly
2450 2455 2460

Asp Phe Thr Ser Leu Val Leu Leu Glu Val Glu Ala His Ala Ala
2465 2470 2475

Ser Thr Phe Ala Glu Arg Ala Ser Arg Leu Gln Ala Gln Leu Trp
2480 2485 2490

Arg Asp Leu Glu His Gly Ser Val Ser Ala Val Gln Leu Ile Arg
2495 2500 2505

Glu Leu Val Arg Thr Gly Arg Arg Ser Pro Gly Ala Ile Met Pro
2510 2515 2520

Val Val Phe Thr Ser Ile Leu Ser Leu Asp Ala Arg Arg Gly Pro
2525 2530 2535

Gln Gly Ser Leu Ser Phe Phe Glu Gly Glu Leu Val Tyr Ser Ile
2540 2545 2550

Ser Gln Thr Pro Gln Val Trp Leu Asp His Gly Val His Glu Glu
2555 2560 2565

Glu Gly Ala Leu Val Leu Ala Trp Asp Ser Val Glu Ala Leu Phe
2570 2575 2580

Pro Pro Gly Met Val Asp Asp Met Phe His Ala Tyr Gln Arg Leu
2585 2590 2595

Leu Gly Ala Leu Ala Glu Glu Glu Gln Ala Trp Glu Gly Glu Leu
2600 2605 2610

Pro Glu Leu Leu Pro Pro Ala Gln Arg Glu Leu Leu Ala Arg Tyr
2615 2620 2625

Asn	Ala	Thr	Gln	Ala	Pro	Arg	Pro	Ser	Gly	Arg	Leu	Glu	Glu	Gly
2630						2635					2640			
Phe	Phe	Thr	Gln	Ala	Arg	Leu	His	Pro	Glu	Leu	Pro	Ala	Leu	Leu
2645						2650					2655			
Ala	Pro	Glu	Arg	Thr	Leu	Ser	Tyr	Gly	Glu	Leu	Ala	Arg	Arg	Ala
2660						2665					2670			
Gln	Ala	Leu	Ala	Ala	Arg	Leu	Arg	Glu	Leu	Glu	Val	Gln	Pro	Gln
2675						2680					2685			
Glu	Leu	Val	Ala	Ile	Ala	Met	His	Lys	Gly	Trp	Glu	Gln	Ala	Thr
2690						2695					2700			
Ala	Val	Leu	Gly	Val	Leu	Gln	Ala	Ala	Ala	Ala	Tyr	Leu	Pro	Leu
2705						2710					2715			
Asp	Pro	Glu	Gln	Pro	Pro	Leu	Arg	Leu	His	Gln	Leu	Leu	Glu	Glu
2720						2725					2730			
Gly	Pro	Ala	Arg	Val	Val	Leu	Thr	Gln	Ser	Ser	Leu	Leu	His	Thr
2735						2740					2745			
Val	Pro	Trp	Pro	Pro	Gly	Val	Gln	Val	Ile	Ala	Val	Asp	Glu	Leu
2750						2755					2760			
Glu	Pro	Ala	Thr	Glu	Ala	Pro	Pro	Leu	Pro	Pro	Arg	Gly	Thr	Pro
2765						2770					2775			
Glu	His	Leu	Ala	Tyr	Val	Ile	Tyr	Thr	Ser	Gly	Ser	Thr	Gly	Lys
2780						2785					2790			
Pro	Lys	Gly	Val	Ala	Ile	Glu	His	Arg	Ala	Ala	Leu	Asn	Thr	Val
2795						2800					2805			
Val	Asp	Leu	Asn	Thr	Arg	Phe	Gly	Val	Gly	Pro	Glu	Asp	Arg	Val
2810						2815					2820			
Leu	Gly	Leu	Ser	Ala	Leu	Thr	Phe	Asp	Leu	Ser	Val	Tyr	Asp	Val
2825						2830					2835			
Leu	Gly	Leu	Leu	Gly	Ala	Gly	Gly	Ala	Leu	Val	Leu	Pro	Ala	Ala
2840						2845					2850			
Glu	Ala	Glu	Lys	Asp	Pro	Ala	His	Trp	Trp	Glu	Arg	Leu	Val	Ala
2855						2860					2865			

Gly Arg Val Thr Val Trp Asn Ser Thr Pro Ala Leu Met Leu Leu
 2870 2875 2880

Leu Val Glu Tyr Ala Glu Gln Arg Gly Leu Lys Leu Pro Ala Ala
 2885 2890 2895

Leu Arg Leu Val Met Leu Ser Gly Asp Trp Ile Pro Val Ala Leu
 2900 2905 2910

Pro Asp Arg Ile Arg Ala Leu Gly Arg Asp Val Gln Val Val Ser
 2915 2920 2925

Leu Gly Gly Ala Thr Glu Ala Ser Ile Trp Ser Ile Ala Tyr Pro
 2930 2935 2940

Ile Gly Gln Val Ala Pro Gln Trp Lys Ser Ile Pro Tyr Gly Met
 2945 2950 2955

Pro Leu Ala Asn Gln Arg Phe His Val Leu Asp Gly Arg Leu Glu
 2960 2965 2970

Ala Arg Pro Trp Trp Val Pro Gly Glu Leu Tyr Ile Gly Gly Glu
 2975 2980 2985

Gly Leu Ala Arg Glu Tyr Trp Arg Asp Glu Pro Leu Thr Ala Thr
 2990 2995 3000

Arg Phe Ile Arg His Pro Arg Thr Gly Glu Arg Leu Tyr Arg Thr
 3005 3010 3015

Gly Asp Gln Gly Arg Met Leu Pro Glu Gly Ser Ile Glu Phe Leu
 3020 3025 3030

Gly Arg Glu Asp Leu Gln Val Lys Val Gln Gly Phe Arg Val Glu
 3035 3040 3045

Leu Gly Glu Ile Glu Ala Ala Leu Ala Gln His Pro Ala Leu Ser
 3050 3055 3060

Ala Ser Val Val Val Ala Arg Gly Glu Pro Arg Gly Val Arg Arg
 3065 3070 3075

Leu Val Ala Tyr Ala Val Pro Arg Ser Gly Gln Thr Pro Ala Ala
 3080 3085 3090

Gly Glu Leu Arg Arg Tyr Leu Ala Glu Arg Leu Pro Ala Tyr Met

3095

3100

3105

Val Pro Ser Ala Phe Val Leu Leu Glu Ser Leu Pro Arg Ser Arg
 3110 3115 3120

Asn Gly Lys Ile Ala Arg Asp Gln Leu Pro Glu Pro Gln Gln Thr
 3125 3130 3135

Gln Gly Leu Ala Ala Gln Ala Ala Ala Ala Asp Pro Leu Val Glu
 3140 3145 3150

Arg Leu Ala Ala Leu Val Lys Glu Ala Leu Arg Leu Glu Arg Val
 3155 3160 3165

Glu Pro Gln Asp Ser Leu Leu Asp Leu Gly Ala Asp Ser Val Ala
 3170 3175 3180

Leu Ile Arg Leu Ile Asn Arg Leu Glu Ala Glu Leu Gln Phe Arg
 3185 3190 3195

Pro Arg Leu Ala Asp Ile Tyr Glu Asn Pro Thr Val Gln Gly Leu
 3200 3205 3210

Ala Thr Leu His Gln Glu Lys Thr Lys Ser Gln Gly Glu Gly Gly
 3215 3220 3225

Ala Pro Arg Leu Thr Ala Pro Arg Ser Thr Leu Leu Pro Ala Glu
 3230 3235 3240

Glu Trp Gly Arg Phe Lys Ala Asn Arg Pro Gly Leu Arg Arg Phe
 3245 3250 3255

Pro Asp Gly Thr Pro Glu Val Ala Leu Pro Gly Ser Gly Leu Ala
 3260 3265 3270

Pro Ala Pro Glu Glu Leu Thr Ala Leu Glu Arg Arg Arg Ser Val
 3275 3280 3285

Arg Thr Tyr Ser Leu Glu Pro Val Ser His Glu Gln Leu Gly Arg
 3290 3295 3300

Leu Leu Ala Pro Leu Arg Glu Trp Glu Val Gln Gly Ser Arg Arg
 3305 3310 3315

Tyr Leu Tyr Ala Ser Ala Gly Gly Leu Tyr Pro Val Gln Leu Tyr
 3320 3325 3330

117

Leu His Leu Lys Pro Gly Arg Ala Arg Gly Leu Glu Pro Gly Thr
 3335 3340 3345

Trp Tyr Tyr Asp Pro Ser Thr His Arg Leu Val Leu Leu Ser Ala
3350 3355 3360

Gly Ala Gly Leu Asp Arg Arg Ile His Asp Pro His Gln Asn Gln
3365 3370 3375

Ala Ile Phe Asp Ser Ala Ala Phe Ser Leu Phe Leu Ile Ala Arg
3380 3385 3390

Met Gly Ala Val Glu Pro Val Tyr Ala Glu His Ala Leu His Phe
3395 3400 3405

Ala Thr Leu Glu Ala Gly Leu Met Thr Gln Leu Leu Asp Leu Gly
3410 3415 3420

Ala Ala Pro Ser Gly Leu Gly Leu Cys His Ile Gly Asp Leu Asp
3425 3430 3435

Phe Ala Gln Ala Arg Gly Leu Phe His Leu Glu Glu Glu His Val
3440 3445 3450

Leu Leu His Ser Leu Val Gly Gly Val Leu Pro Thr Arg Gly Gln
3455 3460 3465

Glu Ala Ala Ser Val Pro Ala Glu Gly Gly Thr Glu Ala Arg Gln
3470 3475 3480

Leu Ala Gln Leu Leu Gln Gln Val Lys Thr Leu Thr Pro Glu Ala
3485 3490 3495

Ala Arg Ala Leu Leu Glu Ala Arg Arg Gly Ser Lys Gly Arg Pro
3500 3505 3510

His Glu
3515

<210> 13

<211> 2837

<212> PRT

<213> Angiococcus disciformis

<400> 13

Met Ser His Asp Thr Gln Leu Glu Gly Ala Val Ala Val Val Ala Met
 1 5 10 15
 Ala Gly Arg Phe Pro Gln Ala Pro Thr Leu Glu Asp Tyr Trp Arg Ile
 20 25 30
 Ile His Glu Gly Val Asp Ala Phe Thr Gln Leu Thr Asp Glu Gln Leu
 35 40 45
 Leu Ala Ser Gly Val Gly Pro Ser Leu Leu Arg Gln Pro Gly Tyr Val
 50 55 60
 Arg Arg Ala Ala Leu Leu Asp Arg Met Asp Gln Phe Asp Ala Ala Phe
 65 70 75 80
 Phe Gly Phe Ser Pro Arg Glu Ala Glu Val Leu Asp Pro Gln His Arg
 85 90 95
 Leu Phe Leu Glu Cys Ala His Glu Ala Leu Glu Arg Ala Gly His Gly
 100 105 110
 Ser Glu Arg Ala Arg Gly Arg Val Gly Val Phe Ala Ser Ala Ser Leu
 115 120 125
 Asn Ser Tyr Tyr Leu His Ser Leu His Gly Asn Ala Arg Leu Arg Glu
 130 135 140
 Val Leu Gly Asp Phe Gln Leu Ala Ile Ala Asn Asp Lys Asp Phe Leu
 145 150 155 160
 Pro Thr Arg Val Ser Tyr Lys Leu Gly Leu Arg Gly Pro Ser Val Ala
 165 170 175
 Val Gln Thr Ala Cys Ser Ser Ser Leu Val Ala Val His Leu Ala Cys
 180 185 190
 Gln Ser Leu Leu Asn Gly Glu Cys Ala Leu Ala Leu Ala Gly Gly Ser
 195 200 205
 Ser Leu Ser Val Pro Gln Ala Gln Gly Tyr Leu Tyr Gln Glu Gly Gly
 210 215 220
 Ile Ala Ser Pro Asp Gly Tyr Cys Arg Pro Phe Asp Ala Ala Ala Ala
 225 230 235 240
 Gly Thr Asn Arg Gly Asn Gly Val Gly Val Val Leu Leu Lys Leu Leu
 245 250 255

Glu Glu Ala Leu Ala Asp Gly Asp Thr Ile His Ala Val Ile Arg Gly
260 265 270

Ser Ala Val Asn Asn Asp Gly Ala His Lys Ile Gly Tyr Thr Ala Pro
275 280 285

Ser Val Glu Gly Gln Ala Ser Val Ile Ser Glu Ala Leu Glu Val Ala
290 295 300

Ala Val Ser Ala Asp Ser Ile Gly Tyr Val Glu Ala His Gly Thr Ala
305 310 315 320

Thr Ala Leu Gly Asp Pro Ile Glu Val Gln Ala Leu Thr Arg Ala Phe
325 330 335

Arg Lys His Thr Glu Arg Arg Gly Tyr Cys Ala Leu Gly Ser Val Lys
340 345 350

Ala Asn Ile Gly His Leu Asp Ala Ala Ala Gly Ile Ala Ser Leu Ile
355 360 365

Lys Ala Val Leu Val Leu Glu Arg Arg Gln Leu Pro Pro Cys Pro His
370 375 380

Phe Thr Ser Pro Asn Pro Arg Ile Asp Phe Glu Arg Ser Pro Phe Tyr
385 390 395 400

Val Ser Gly Arg Gly Gln Pro Trp Glu Pro Val Asp Gly Pro Arg Arg
405 410 415

Ala Gly Val Ser Ser Phe Gly Ile Gly Gly Thr Asn Val His Val Val
420 425 430

Leu Glu Glu Ala Pro Pro Arg Pro Ala Pro Arg Arg Ala Ala Arg Pro
435 440 445

Tyr Gln Leu Leu Pro Leu Ser Ala Arg Thr Glu Arg Ala Arg Glu Glu
450 455 460

Val Glu Ala Arg Leu Arg Glu His Leu Arg Gln Gln Pro Gln Glu Pro
465 470 475 480

Leu Glu Glu Val Ala His Thr Leu Gln Val Gly Arg Arg His Met Ala
485 490 495

Trp Arg Thr Ala Leu Val Ser Ser Ser Pro Ala Gln Ala Val Glu Leu

500

505

510

Leu Glu Arg Arg Arg Pro Glu Ala Leu Leu Arg Gly Gln Ser Ala Ala
515 520 525

Gln Ala Arg Ser Val Ala Phe Leu Leu Pro Gly Gln Gly Ser Gln Tyr
530 535 540

Val Gly Met Gly Ala Ala Leu His Glu Ser Glu Ala Pro Phe Arg Glu
545 550 555 560

Gln Val Asp Leu Cys Ala Gly Lys Leu Gln Gly Val Leu Gly Leu Asp
565 570 575

Val Arg Arg Leu Leu Tyr Pro Gly Pro Arg Glu Arg Glu Trp Ala Gln
580 585 590

Glu Arg Leu Arg Glu Thr Arg Val Cys Gln Pro Val Leu Phe Thr Val
595 600 605

Glu Tyr Ala Leu Ala Arg Leu Leu Glu Ala Trp Gly Ile Arg Pro Thr
610 615 620

Ala Leu Leu Gly His Ser Leu Gly Glu Tyr Val Ala Ala Cys Leu Ala
625 630 635 640

Gly Val Phe Ser Leu Glu Glu Ala Leu Glu Val Val Ala Ala Arg Gly
645 650 655

Gln Leu Met Gly Ser Met Pro Pro Gly Gly Met Leu Ala Val Gly Leu
660 665 670

Pro Ala Gln Glu Val Glu Pro Leu Leu Pro Ala Ala Leu Ala Leu Ala
675 680 685

Ala His Asn Ser Pro Gln Ala Cys Val Val Ala Gly Pro Thr Glu Pro
690 695 700

Leu Ala Glu Leu Arg Ala Arg Leu Glu Gln Arg Gly Ile Ala Cys Thr
705 710 715 720

Pro Leu Ala Val Ser His Ala Phe His Ser Pro Met Met Glu Pro Ala
725 730 735

Val Ala Pro Phe Val Ala Arg Leu Arg Arg Met His Leu Arg Pro Pro
740 745 750

Ser Leu Pro Phe Ile Ser Asn Val Thr Gly Thr Trp Ile Glu Ala Glu
755 760 765

Glu Ala Thr Ser Pro Glu Tyr Trp Gly Arg His Leu Leu Gln Pro Val
770 775 780

Arg Phe Ala Gln Gly Leu Glu Arg Leu Cys Glu Gly Val Gln Pro Arg
785 790 795 800

Leu Leu Leu Glu Val Gly Pro Gly His Thr Leu Gly Arg Leu Ala Ala
805 810 815

Arg Gln Ser Ser Gly Pro Val Gln Val Val Ser Thr Leu Gly Ser Ser
820 825 830

Arg Glu Glu Ser Ser Glu Val Glu Arg Leu Leu Thr Ala Ile Gly Arg
835 840 845

Leu Trp Val Glu Gly Ala Glu Val Asp Trp Ala Gly Leu His Arg Gly
850 855 860

Glu Arg Arg Arg Arg Ala Val Leu Pro Thr Tyr Pro Phe Glu His Gln
865 870 875 880

Arg Tyr Trp Val Glu Thr Ala Pro Gln Pro Ser Pro Pro Glu Arg Pro
885 890 895

Gly Thr Pro Ala Glu Ser Pro Val Thr Ser Ser Phe Tyr Val Pro Gly
900 905 910

Trp Ser Arg Ala Ala Leu Pro Ser Ala Ala Pro Ser Arg Arg Val Gly
915 920 925

Arg Leu Leu Leu Leu Ala Glu Ala His Gly Trp Ser Gln Gly Leu Ala
930 935 940

Glu Arg Leu Arg Ala Arg Gly His Ser Val Thr Leu Val Glu Pro Gly
945 950 955 960

Glu Arg Leu Glu Arg Leu Thr Pro Glu His Trp Arg Leu Pro Pro Gly
965 970 975

Arg Arg Glu Asp Phe Gln Arg Leu Leu Glu Asp Ser Gly Glu Leu Pro
980 985 990

Ser Arg Val Leu His Leu Trp Leu Leu Gly Ser Ser Gln Ser Pro Gln
995 1000 1005

Glu Arg 1010	Gly Phe Tyr Thr	Leu 1015	Leu Ala Leu Ala	Gln 1020	Ala Leu Gly
Ala His 1025	Gly Thr Arg Pro	Ala 1030	Val Glu Leu Thr	Val 1035	Val Thr Asp
Gln Leu 1040	Tyr Ala Val Ala	Glu 1045	Gly Glu Pro Ile	Gln 1050	Pro Leu Lys
Ala Leu 1055	Leu Gln Gly Pro	Ala 1060	Ser Val Leu Pro	Gln 1065	Glu Leu Pro
Gly Cys 1070	Thr Cys Arg Leu	Val 1075	Asp Met Ala Leu	Pro 1080	Pro Gly Gly
Val Ala 1085	Glu Glu Leu Leu	Glu 1090	Arg Leu Val Ala	Glu 1095	Val Glu Ser
Thr Ala 1100	Ser Glu Arg Ser	Val 1105	Ala Tyr Arg Gly	Ala 1110	Ala Arg Trp
Val Arg 1115	Glu Phe Val Pro	Val 1120	Pro Leu Pro Pro	Pro 1125	Glu Pro Ser
Gln Leu 1130	Pro Leu Arg Gln	Arg 1135	Gly Val Tyr Leu	Ile 1140	Val Gly Gly
Leu Gly 1145	Gly Val Gly Leu	Ala 1150	Leu Ala Glu His	Leu 1155	Ala Arg Arg
Val Ser 1160	Ala Arg Leu Val	Leu 1165	Thr Gly Arg Ser	Pro 1170	Thr Pro Pro
Arg Glu 1175	Ser Trp Ser Ala	Trp 1180	Leu Gly Thr Pro	Thr 1185	Arg Leu Arg
Leu Ser 1190	Gln Glu Leu Glu	Trp 1195	Leu Arg Gly Ala	Ala 1200	Glu Gln Ile
Glu His 1205	Gln Arg Pro Leu	His 1210	Ser Leu Ala Asp	Ser 1215	Pro Gly Leu
Glu Glu 1220	Ser Leu Arg Leu	Leu 1225	Cys Ala Ser Tyr	Leu 1230	Tyr His Phe
Leu Phe 1235	Pro Leu Gln Ala	Pro 1240	Leu Arg Leu Gly	Glu 1245	Pro Arg Ala

Met Gln Ala Leu Arg Glu Arg Leu Gly Leu Gln Pro Gly Phe Glu
 1250 1255 1260
 Arg Leu Phe Ser Phe Met Ile Gly Thr Leu Glu Gln Ala Lys Leu
 1265 1270 1275
 Ile Arg Val Gly Glu Gly Met Leu Glu Gly Gln Val Glu Pro Ala
 1280 1285 1290
 His Val Pro Thr Pro Arg Ala Leu His Glu Arg Leu Leu Glu Gly
 1295 1300 1305
 Tyr Pro Glu Val Ser Gly Leu Leu Glu Leu Leu Glu His Cys Val
 1310 1315 1320
 Arg His Tyr Pro Glu Val Leu Arg Gly Ser Leu Pro Ala Leu Ser
 1325 1330 1335
 Val Leu Tyr Pro Ser Gly Arg Ser Glu Glu Asp Ser Ala Arg Ser
 1340 1345 1350
 Gly Val Glu Trp Ser Ser Met Gly Gln Cys Val Ala Leu Leu Arg
 1355 1360 1365
 His Phe Leu Ala Arg His Ala Glu Arg Thr Gln Gly Arg Thr Leu
 1370 1375 1380
 Arg Ile Leu Glu Val Gly Gly Gly Ser Gly Val Val Leu Gln Ala
 1385 1390 1395
 Leu Leu Pro Leu Leu Arg His His Pro Val Glu Tyr His Phe Thr
 1400 1405 1410
 Asp Ile Gly Pro Ser Phe Val Arg Ala Met Glu Gly Val Gly Arg
 1415 1420 1425
 Arg Gln Gly Leu Thr Phe Leu His Thr Ser Val Leu Asp Ile Ser
 1430 1435 1440
 Arg Pro Pro Pro Glu Gln Gly His Pro Pro Gly Ser Tyr Asp Leu
 1445 1450 1455
 Val Ile Ala Leu Asn Val Val His Ala Thr Pro Arg Val Pro Gln
 1460 1465 1470
 Ser Leu Ala His Leu Glu Ser Leu Leu Val Pro Gly Gly His Leu

1475		1480		1485
Cys Leu Val Glu Thr Val	Lys Gln Gln Pro Trp Val	Asp Met Ile		
1490	1495	1500		
Trp Gly Leu Ala Glu Gly	Trp Trp Ser Tyr Glu Asp	Glu Leu Arg		
1505	1510	1515		
Thr Arg Ser Pro Leu Leu	Glu Val Gly Asp Trp Glu	His Ala Leu		
1520	1525	1530		
Arg Asp Val Gly Phe Ala	Glu Val Glu Val Leu Pro	Ala Ala Val		
1535	1540	1545		
Glu Gln Arg Ser Arg Trp	Asp Asn Val Leu Leu Ile	Ala Gln Arg		
1550	1555	1560		
Pro Gly Glu Ser Gly Leu	Gln Pro Val Gly Gly Arg	Ala Ala Met		
1565	1570	1575		
Gln Glu Arg Ile Arg Arg	Leu Arg Ala Ile Glu Glu	Ala Gly Gly		
1580	1585	1590		
Glu Val Leu Pro Leu Val	Ala Asp Val Thr Asp Arg	Glu Arg Met		
1595	1600	1605		
Gly Glu Val Leu Ala Glu	Val Lys Arg Arg His Gly	Ala Leu His		
1610	1615	1620		
Gly Val Ile His Ala Ala	Leu Val Leu Glu Asp Gly	Leu Met Gln		
1625	1630	1635		
Leu Lys Thr Arg Glu Ser	Ala Gly Arg Val Leu Ala	Ser Lys Val		
1640	1645	1650		
Glu Gly Thr Leu Val Leu	Asp Glu Leu Leu Arg Asp	Glu Pro Leu		
1655	1660	1665		
Asp Phe Phe Val Leu Cys	Ser Ser Leu Ser Ala Leu	Leu Gly Ala		
1670	1675	1680		
Leu Gly Gln Ala Asp Tyr	Ala Ala Ala Ser Ala Phe	Leu Asp Ala		
1685	1690	1695		
Tyr Ala His Ser Gln Arg	Gly Arg Thr Asp Arg Arg	Thr Ile Ser		
1700	1705	1710		

Met Asp Trp Asp Arg Trp Leu Glu Val Gly Ala Ala Met Arg Leu 1715 1720 1725
Gly Leu Gly Leu Ala Ala Gly Ala Leu Gly Leu Gln Arg Thr Ala 1730 1735 1740
Pro Gly Glu Tyr Thr Val Arg Trp Gln Ala Glu Arg Cys Trp Trp 1745 1750 1755
Leu Asp Glu His Arg Leu Glu Gly Arg Ala Thr Leu Pro Gly Val 1760 1765 1770
Ala Tyr Leu Glu Leu Val Arg Ala Ala Leu Val Gln Glu Leu Gly 1775 1780 1785
Glu Ala Pro Val Glu Leu Glu Gln Leu Val Leu Leu Ser Leu Leu 1790 1795 1800
Glu Ala Pro Ala Gly Glu Glu Val Glu Val Arg Phe His Leu Arg 1805 1810 1815
Pro Glu Asp Glu Gly Tyr Ala Leu Glu Ile Arg Ser Arg Ala Gly 1820 1825 1830
Gly Leu Ala Asn Gly Gly Trp Arg Ser His Ala Met Gly Arg Val 1835 1840 1845
Arg Met Leu Pro Arg Gly Ser Ala Arg Pro Thr His Pro Leu Arg 1850 1855 1860
Glu Leu Glu Glu Arg Leu Gly Leu Thr Gln Ala Pro Arg Glu His 1865 1870 1875
Glu Pro Ala Gln Gly Pro Arg Glu Leu Ala Pro Ala Leu Gly Arg 1880 1885 1890
Arg Trp Ser Ser Leu Ser Trp His Arg Ala Trp Lys Gly Glu Glu 1895 1900 1905
Gly Leu Ala Leu Ile Glu Leu Pro Glu Glu Leu Ala Glu Asp Leu 1910 1915 1920
Arg Gln Trp Pro Leu His Pro Ala Leu Leu Asp Ala Ala Thr Gly 1925 1930 1935
Phe Ala Pro Leu Pro Pro Gly Ala Trp Leu Pro Leu Ser Tyr Gly 1940 1945 1950

Asn Thr 1955	Arg Ile His Gly 1960	Pro Leu Pro Arg Gln Leu 1965	Tyr Ser His
Ile Arg 1970	Arg Leu Glu Pro 1975	Ala Ser Ala Gln Ala Gly 1980	Val Val Arg
Leu Glu 1985	Val Arg Leu Met 1990	Asp Gly Glu Gly Arg 1995	Leu Leu Cys
Val Glu 2000	Glu Phe Val Leu 2005	Arg Arg Val Glu Val 2010	Asp Ser Leu Ala
Arg Pro 2015	Gln Pro Ala Gly 2020	Lys Ala Pro Arg Gln 2025	Ala Val Leu Pro
Arg Pro 2030	Gly Ala Leu Asp 2035	Ser Leu Arg Leu Gln 2040	Pro Leu Glu Arg
Leu Pro 2045	Pro Gln Glu Gly 2050	Glu Val Glu Val Gln 2055	Val Leu Ala Ala
Gly Leu 2060	Asn Phe Lys Asp 2065	Ala Leu Leu Ala Leu 2070	Gly Ala Leu Pro
Val Glu 2075	Leu Ala Asn Gly 2080	Ala Pro Leu Ala Leu 2085	Gly Val Glu Cys
Ala Gly 2090	Ile Ile Ser Ala 2095	Val Gly Pro Gly Val 2100	Arg Gly Leu Arg
Val Gly 2105	Glu Ala Val Val 2110	Ala Ala Ala Gly 2115	Ala Phe Ala Ser
His Val 2120	Arg Val Pro Gln 2125	Glu Gln Val Phe Pro 2130	Lys Pro Ala Gly
Leu Ser 2135	Phe Glu Gln Ala 2140	Ala Met Val Pro Val 2145	Thr Leu Phe Thr
Ala Trp 2150	Tyr Ala Leu Glu 2155	Glu Leu Ala Arg Leu 2160	Arg Ala Gly Glu
His Val 2165	Leu Ile His Ala 2170	Ala Ala Thr Gly Val 2175	Gly Leu Ala Gly
Val Lys 2180	Leu Ala Leu Arg 2185	Arg Gly Ala Thr Val 2190	Tyr Ala Thr Ala

Gly Ser Glu Pro Lys Arg Glu Leu Leu Arg Ser Leu Gly Val Thr
 2195 2200 2205
 Leu Ala Met Asp Ser Arg Ala Pro Gly Phe Asp Asp Gln Leu Leu
 2210 2215 2220
 Gln His Thr Gln Gly Arg Gly Val Asp Val Val Leu Asn Ser Leu
 2225 2230 2235
 Ser Gly Glu Phe Leu Ser Arg Ser Leu Gly Val Leu Ala Arg His
 2240 2245 2250
 Gly Arg Phe Val Glu Leu Gly Val Arg Asp Met Leu Ser Gly Gly
 2255 2260 2265
 Thr Leu Pro Leu Ala Pro Phe Glu Arg Gly Leu Thr Phe Leu Ala
 2270 2275 2280
 Ala Gln Ile Asp Pro Gly Met Lys Gly Tyr Arg Glu Leu Met Gly
 2285 2290 2295
 Glu Ala Leu Arg Gln Ile Glu Arg Gly Glu Leu Glu Pro Leu Pro
 2300 2305 2310
 Tyr Thr Ala Trp Pro Leu Glu Arg Val Ala Glu Ala Leu Gln Leu
 2315 2320 2325
 Val Ser Lys Gly Arg His Val Gly Lys Val Val Leu Thr Pro Glu
 2330 2335 2340
 Glu Pro Leu Ser Arg Pro Pro Ala Leu Pro Ser Ser Ser Arg Pro
 2345 2350 2355
 Val Gln Pro Arg Ala Gly Val Ala Ile Val Gly Leu Arg Ser Ala
 2360 2365 2370
 Glu Gly Cys Glu Ala Phe Glu Arg Ile Leu Ala Ala Gly Leu Pro
 2375 2380 2385
 Gln Val Ala Val Ser Thr Arg Glu Leu Arg Ala Arg Met Ala Glu
 2390 2395 2400
 Ile Glu Arg Leu Arg Val Ser Ser Trp Glu Pro Ala Leu Pro Ser
 2405 2410 2415
 Val Pro Arg Gly Ser Pro Ala Arg Thr Glu Arg Ser Lys Pro Tyr

2420		2425		2430
Val Ala Pro Arg Thr Glu Arg Glu Arg Ala Leu Ala Ala Leu Trp				
2435		2440		2445
Gln Asp Leu Leu Gly Val Ser Glu Val Gly Leu Asp Asp Asp Phe				
2450		2455		2460
Phe Glu Leu Lys Gly Asp Ser Leu Leu Ala Ile Gln Leu Met Gly				
2465		2470		2475
Arg Leu Arg Lys Glu Leu Glu Val Glu Leu Pro Leu Ser Ser Phe				
2480		2485		2490
Leu Ala Arg Pro Thr Leu Arg Thr Leu Leu Ser Ala Met Pro Ala				
2495		2500		2505
Pro Ala Thr Asp Glu Pro Ala Ser Ala Thr Gly Ala Ile Pro Ala				
2510		2515		2520
Ala Ala Pro Gln Pro Ser Pro Ala Val Ser Ser Glu Pro Ala Arg				
2525		2530		2535
Lys Trp Arg His Leu Val Pro Ile Gln Pro Gln Gly Glu Ser Pro				
2540		2545		2550
Pro Phe Phe Trp Ala Ala Pro Leu Met Gly Thr Val Phe Pro Tyr				
2555		2560		2565
Phe Thr Leu Ala Arg Cys Leu Ala Pro Thr His Pro Phe Tyr Ala				
2570		2575		2580
Leu Gln Pro Pro Gly Leu Gln Glu Gly Gln Glu Pro Met Asp Lys				
2585		2590		2595
Val Glu Glu Leu Ala Ala Leu Tyr Val Arg Glu Leu Arg Glu Leu				
2600		2605		2610
Gln Pro Arg Gly Pro Tyr Arg Leu Gly Gly Trp Ser Phe Gly Cys				
2615		2620		2625
Val Val Ala Tyr Glu Val Ala Leu Gln Leu Glu Ala Ala Gly Glu				
2630		2635		2640
Gln Val Ala Leu Leu Ser Leu Leu Asp Phe Pro Ala Pro Ser Gly				
2645		2650		2655

Gln Arg Ala Ser Leu Trp Ala Thr Ala Arg Phe Phe Ser Gly Ser
 2660 2665 2670

Val Leu Arg Gly Leu Ala Pro Tyr Thr Leu Asp Tyr Leu Tyr Leu
 2675 2680 2685

Ala Ala Arg Ala Pro Leu Asn Pro Asp Ala Ala Ser Arg Leu Ser
 2690 2695 2700

Ser Ser Leu Thr Arg Gly Trp Glu Arg Leu Arg Lys Gly Gly Leu
 2705 2710 2715

Val Gln Glu Leu Leu Asp His Ala Ala Met Ala Arg Val Val Pro
 2720 2725 2730

Glu Asp Ser Arg Leu Leu Leu Leu Arg Glu Pro Gly Ile Ala Pro
 2735 2740 2745

Met Leu Arg Leu Thr Arg Ala His Gln Arg Ala Met Leu Ser Tyr
 2750 2755 2760

Arg Pro Arg Gly Arg Leu Arg Gln Arg Ile Val Leu Phe Arg Thr
 2765 2770 2775

Ser Glu Gln Ala Trp Ser Phe Ala Arg Asp Leu Gly Trp Gly Ala
 2780 2785 2790

Leu Ser Ala Ser Gly Val Asp Val Arg Glu Ala Pro Gly Asp His
 2795 2800 2805

Met Thr Leu Leu Arg Pro Pro His Val Glu Arg Val Ala Glu Gln
 2810 2815 2820

Leu Arg Ala Leu Leu Ser Ser Ala Pro Leu Pro Thr Arg Arg
 2825 2830 2835

<210> 14

<211> 385

<212> PRT

<213> Angiococcus disciformis

<400> 14

Met Lys Phe Pro Pro Ala Ser Leu Cys Phe Val Leu Pro Met Leu Ser
 1 5 10 15

Ala Thr Ser Val Gln Ala Gln Ala Pro Glu Thr Trp Thr Glu His Trp
20 25 30

Phe Glu His Asn Gln Thr Val Ser Arg Val Tyr Gln Asp Asn Asp Val
35 40 45

Ala Val Tyr Phe Asp Ser Ala Val Asn Arg Ser Ile Thr Trp Pro Asn
50 55 60

Thr Tyr Leu Gly Asp Val Trp Arg Tyr Thr Lys Arg Thr Tyr Gly His
65 70 75 80

Phe Gly Thr Asp Ser Gln Leu Tyr Thr Ile Phe His Ala Gly Lys Tyr
85 90 95

Ser Gly Gly His Pro Ser Thr Tyr Phe Asp Ala Ser His Asp Arg Arg
100 105 110

Asn Val Ile Asp Val Gly Ser Ser Ser Pro Thr Ala Trp Thr Ala Gly
115 120 125

Thr Gly Asn Asp Leu Asp Ile Val Thr His Glu Val Ala His Ile Val
130 135 140

Glu Leu Ala Ser Lys Gly Val His Asp Ser Pro Ala Phe Gly Leu Trp
145 150 155 160

Arg Asp Ser Lys Trp Ala Glu Ile Phe Asn Tyr Asp Val Tyr Leu Ala
165 170 175

Leu Gly Arg Thr Ser Asp Ala Asn Arg Trp Tyr Asn Leu Met Val Asn
180 185 190

Thr Thr Asp Ser Phe Pro Arg Ala Asn Thr His Trp Phe Arg Asp Trp
195 200 205

Phe Tyr Pro Ile Tyr Lys Asn His Gly Gly Ser Ser Val Leu Asn Arg
210 215 220

Tyr Phe Val Leu Leu Ala Gln Tyr Leu Pro Lys Asn Gly Ser Asn Tyr
225 230 235 240

Ala Arg Ser Leu Asn Trp Gly Glu Phe Ile His Phe Trp Ser Gly Ala
245 250 255

Ala Gly Val Asn Leu Lys Thr Leu Ala Thr Ser Ala Phe Gly Trp Pro
260 265 270

Ala Glu Trp Glu Ala Gln Phe Thr Gln Ala Gln Arg Asp Phe Pro Phe
 275 280 285

Thr Tyr Thr Pro Pro Gly Pro Thr Ala Val Thr Val Phe Gln Asp Gln
 290 295 300

Asn Tyr Gly Gly Tyr Gly Met Ala Leu Pro Val Gly Arg Tyr Thr Leu
 305 310 315 320

Ser Ala Leu His Ala Trp Gly Val Arg Asn Asp Asp Ile Thr Ser Leu
 325 330 335

Lys Val Ala Ser Gly Tyr Lys Val Thr Leu Tyr Glu His Asp Asn Phe
 340 345 350

Ser Gly Ala Ser Leu Thr Lys Thr Ala Asp Asp Ala Ser Leu Val Asp
 355 360 365

Asp Ser Trp Asn Asp Lys Met Ser Ser Leu Ile Val Ser Ala Ser Gly
 370 375 380

Met
 385

<210> 15

<211> 252

<212> PRT

<213> *Angiococcus disciformis*

<400> 15

Met Phe Asp Ala Thr Arg Tyr Leu Glu Arg Ile Gly Val Ser Ala Gly
 1 5 10 15

Thr Pro Leu Ala Ala Leu His Arg Ala His Leu Glu Ala Val Pro Phe
 20 25 30

Glu Asn Leu Asp Ile His Leu Lys Arg Pro Ile Arg Leu Asp Glu Ala
 35 40 45

Ala Phe Phe Glu Lys Val Val Leu Gln Arg Arg Gly Gly Phe Cys Tyr
 50 55 60

Glu Leu Asn Gly Leu Phe Ala Arg Leu Leu Arg Thr Leu Gly Tyr Arg

132

65

70

75

80

Val Thr Leu Leu Ser Ala Arg Val Ala Ser Arg Pro Asp Gly Ser Ala
85 90 95

Tyr Gly Pro Asp Phe Asp His Leu Ala Leu Leu Val Glu Asp Ala Ser
100 105 110

Gly Arg Trp Leu Ala Asp Val Gly Phe Gly Glu Cys Phe Leu Glu Pro
115 120 125

Leu Arg Leu Asp Glu Arg Gly Val Gln Thr Gln Asp Gly Arg Gly His
130 135 140

Arg Leu Val Glu Asp Ala Glu Gly Leu Val Val Trp Arg Glu Ala Ala
145 150 155 160

Ser Gly Asp Trp Lys Ala Gln Tyr Val Val Ser Leu Ile Pro Arg Glu
165 170 175

Leu Gly Asp Phe Ala Ala Met Cys His His Gln Gln Thr Ser Pro Glu
180 185 190

Ser His Phe Thr Gln Arg Arg Leu Cys Thr Arg Thr Thr Pro Asp Gly
195 200 205

Arg Ile Thr Leu Lys Glu Gly Ala Leu Val Val Thr Ser Gly Gly Ala
210 215 220

Arg His Glu Gln Arg Leu Pro Asp Glu Ala Ala Trp Arg Ala Ala Leu
225 230 235 240

Thr Arg His Phe Gly Ile Thr Leu Pro Ser Gly Gly
245 250

<210> 16

<211> 308

<212> PRT

<213> Angiococcus disciformis

<400> 16

Met Thr Thr Ser Lys Val Arg Ser Gly Ala Thr Leu Gly Gly Val Thr
1 5 10 15

133

Ala Arg Phe Ser Thr Gly Glu Gln Pro Ala Pro Lys Lys Tyr Val Asp
20 25 30

Tyr Thr Val Lys Ser Gly Asp Thr Phe Ser Gly Ile Ala Ala Lys His
35 40 45

Glu Leu Ser Leu Ala Ala Leu Lys Lys Leu Asn Pro Gln Val Glu Asn
50 55 60

Ile Asn Val Ile His Pro Gly Gln Lys Leu His Val Lys Val Val Asn
65 70 75 80

Ala Pro Ala Pro Glu Pro Lys Pro Thr Pro Lys Pro Pro Thr Asn Lys
85 90 95

Asn Thr Ser Val Leu Pro Lys Gly Ile Pro Asn Thr Glu Gly Met Ser
100 105 110

Gln Ala Lys Glu Tyr Ala Leu Tyr Ser Lys Tyr Val Glu Lys Tyr Gly
115 120 125

Asp Ala Lys Ala Lys Gln Asp Leu Ala Ala Gly Lys Arg Val Ile Val
130 135 140

Gly Leu Arg Val Asn Thr Pro Phe Thr Arg Asp Arg Pro Ser Gly Gly
145 150 155 160

Ser Tyr Asp Asp Arg Leu Val Val Met Trp Lys Asp Ser Ser Gly Lys
165 170 175

Pro His Val Gln Glu Phe Lys Ala Asn Thr Glu Pro Asn Arg Arg Trp
180 185 190

Ala Asp Asp Pro Ser Gln Ser Thr Lys Pro Val Gly Arg Leu Val Gly
195 200 205

Asn Lys Thr Tyr His Tyr Arg Lys Ser Phe Asn Gly Asn Phe Gly Gly
210 215 220

Asn Ile Leu Ala Pro Asp Leu Arg Tyr Gly Asn Pro Thr Val Arg Arg
225 230 235 240

Asp Thr Asn Arg Asp His Arg Ile Asn Ser Lys Asp Asp Val Phe Ser
245 250 255

Gly Asp Trp Gly Gly Gln Gly Tyr Tyr Phe His Arg Gly Gly Thr Thr
260 265 270

Asp Thr Tyr Ser Ala Gly Cys Gln Thr Met Asp Gln Gly Arg Phe Asn
 275 280 285

Asn Phe Trp Ala Ala Leu Gly Ser Gln Ser Glu Phe Ser Tyr Val Leu
 290 295 300

Ala Gln Val Gly
 305

<210> 17

<211> 293

<212> PRT

<213> Angiococcus disciformis

<400> 17

Met Thr Arg Val Glu Gln Leu Arg Asp Leu Glu Thr Ala Lys Gln Gly
 1 5 10 15

Ala Ala Leu Leu Glu Ala Glu Asn Ala Arg Leu His Gln Arg Leu Glu
 20 25 30

Ala Leu Val Ala Glu Asn Ala Arg Leu Lys Gly Glu Asp Ala Gln Ala
 35 40 45

His Leu His His Ala Asp Glu Thr His Trp Leu Leu Leu Asp Lys Gly
 50 55 60

Pro Gly Thr Lys Trp Tyr Ala Trp Thr Val Ala Ser Pro Asp Thr Val
 65 70 75 80

Phe His Arg Ile Tyr Pro Ser Arg Ser Gly Ala Thr Ala Arg Thr Val
 85 90 95

Leu Gly Asp Tyr Glu Gly Val Val Leu Val Asp Gly Tyr Ala Ala Tyr
 100 105 110

Gln Thr Ala Thr Lys Ser Ser Ala Asp Gly Pro Cys Pro Ala Thr Leu
 115 120 125

Ala Phe Cys Trp Ala His Val Arg Arg Lys Phe Phe Glu Ala Gln Gln
 130 135 140

Phe Ala Pro Ala Cys Lys Glu Val Leu Asp Leu Ile Gly Glu Leu Tyr
 145 150 155 160

Ala Val Glu Ala Asp Leu Pro Gly Trp Tyr Ala Leu Glu Gly Glu Glu
 165 170 175

Arg Gln Ala Ala Leu Ala His Arg Leu Ala Val Arg Gln Gln Lys Ser
 180 185 190

Ala Pro Leu Thr Glu Arg Ile Arg Asp Trp Ala His Ala Gln Arg Ala
 195 200 205

Leu Pro Gly Ser Ala Phe Arg Lys Ala Leu Glu Tyr Met Leu Asn Leu
 210 215 220

Trp Ala Gly Leu Thr Val Phe Leu Thr Gln Pro Gln Val Pro Leu Asp
 225 230 235 240

Asn Asn His Val Glu Arg Gln Leu Arg Asp Met Val Ile Gly Arg Lys
 245 250 255

Asn Leu Lys Thr Ala Arg Leu Arg Gly Glu Ala Pro Gly His Tyr Leu
 260 265 270

Arg Arg Pro Ala Leu Ala Ala Ile Glu Asn Pro Gly Thr Val Thr Leu
 275 280 285

Pro Lys Ala Gln Ser
 290

<210> 18

<211> 82868

<212> DNA

<213> *Angiococcus disciformis*

<220>

<221> CDS

<222> (72)..(1148)

<223> ORF24

<220>

<221> CDS

<222> (1550)..(2812)

<223> ORF23

<220>

<221> CDS

<222> (7660)..(8214)

<223> ORF19

<220>

<221> CDS

<222> (10083)..(11138)

<223> ORF18

<220>

<221> CDS

<222> (11229)..(12284)

<223> ORF17

<220>

<221> CDS

<222> (47447)..(48664)

<223> ORF1

<220>

<221> CDS

<222> (48674)..(49795)

<223> tubZ

<220>

<221> CDS

<222> (49813)..(50472)

<223> ORF2

<220>

<221> CDS

<222> (61715) .. (64060)

<223> ORF7

<220>

<221> CDS

<222> (68036) .. (69034)

<223> ORF10

<220>

<221> CDS

<222> (70676) .. (72319)

<223> ORF12

<220>

<221> CDS

<222> (73312) .. (75552)

<223> ORF13

<220>

<221> CDS

<222> (78163) .. (79416)

<223> ORF15

<220>

<221> CDS

<222> (79561) .. (82866)

<223> ORF16

<400>	18															60	
attacgaatt	cgcggccgca	attaaccctc	actaaaggga	tcaatcgaaa	gcacctcgag											60	
attcgagggtg	t	gtg	ggg	cgt	gga	ctt	cag	cca	ctt	cca	ccg	gat	gct	ccg	110		
	Val	Gly	Arg	Gly	Leu	Gln	Pro	Leu	Pro	Pro	Asp	Ala	Pro		110		
	1	5					10										
cat	cct	cga	ccg	gca	cgg	tgc	gag	ggt	cgc	cgg	tgg	caa	ggg	ctg	gga	158	
His	Pro	Arg	Pro	Ala	Arg	Cys	Glu	Gly	Arg	Arg	Trp	Gln	Gly	Leu	Gly	158	
	15	20					25										
gca	cat	ggt	cag	ccg	cac	cta	cgg	cct	cga	aga	ggt	gaa	tca	ggc	cct	206	
Ala	His	Gly	Gln	Pro	His	Leu	Arg	Pro	Arg	Arg	Gly	Glu	Ser	Gly	Pro	206	
	30	35					40					45					
ggc	cga	tgt	cga	gac	ggg	tcg	cat	cgt	gaa	ggg	act	cat	ccg	ccc	caa	254	
Gly	Arg	Cys	Arg	Asp	Gly	Ser	His	Arg	Glu	Gly	Thr	His	Pro	Pro	Gln	254	
				50	55					60							
tcc	cgt	tcg	tac	ctc	gcg	atg	acg	gcc	gat	tcg	atg	gat	gcc	acc	gtg	302	
Ser	Arg	Ser	Tyr	Leu	Ala	Met	Thr	Ala	Asp	Ser	Met	Asp	Ala	Thr	Val	302	
			65	70					75								
ccc	tcg	cgg	agg	cca	gcc	ctg	ggc	ttc	atc	ggc	ggt	ggg	gtg	atg	gcc	350	
Pro	Ser	Arg	Arg	Pro	Ala	Leu	Gly	Phe	Ile	Gly	Gly	Gly	Val	Met	Ala	350	
		80	85					90									
gag	gcc	ttg	att	gga	gcg	gtc	acc	ggg	gga	ggc	ctg	gcc	gcg	ccc	ggc	398	
Glu	Ala	Leu	Ile	Gly	Ala	Val	Thr	Gly	Gly	Gly	Leu	Ala	Ala	Pro	Gly	398	
	95	100					105										
acg	ctc	ctc	gtc	gga	gag	gtc	gac	gag	gag	cgt	cgc	cgg	gtg	ctc	cgc	446	
Thr	Leu	Leu	Val	Gly	Glu	Val	Asp	Glu	Glu	Arg	Arg	Arg	Val	Leu	Arg	446	
	110	115					120					125					
gag	cag	tat	ggc	gtc	gag	gtg	acc	gcg	gac	aac	ctc	gcg	gcg	gcg	gac	494	
Glu	Gln	Tyr	Gly	Val	Glu	Val	Thr	Ala	Asp	Asn	Leu	Ala	Ala	Ala	Asp	494	
			130	135					140								
gcg	gaa	ctg	gtc	gtg	ctg	gcc	gtg	aag	ccg	caa	cac	ctg	ggt	gtt	gtc	542	
Ala	Glu	Leu	Val	Val	Leu	Ala	Val	Lys	Pro	Gln	His	Leu	Gly	Val	Val	542	
			145	150					155								
ctc	cag	cag	ctg	cgc	ggg	cag	ctc	cgt	gag	gac	acg	att	gtc	gtc	agc	590	
Leu	Gln	Gln	Leu	Arg	Gly	Gln	Leu	Arg	Glu	Asp	Thr	Ile	Val	Val	Ser	590	
		160	165					170									
atc	atc	gcg	ggc	gcg	cgc	ctg	cag	acg	ctg	tgc	gag	ggt	ctc	gaa	cac	638	
Ile	Ile	Ala	Gly	Ala	Arg	Leu	Gln	Thr	Leu	Cys	Glu	Gly	Leu	Glu	His	638	
	175	180					185										
cca	ctc	gtc	gtt	cgc	gcg	atg	ccc	aac	acc	ccc	gcg	cgg	gtc	cgc	aag	686	
Pro	Leu	Val	Val	Arg	Ala	Met	Pro	Asn	Thr	Pro	Ala	Arg	Val	Arg	Lys	686	
	190	195					200					205					
gcg	gcc	acc	ttc	tgg	att	cca	agc	gcg	ggc	ttg	agc	gag	ccg	gcg	ctg	734	
Ala	Ala	Thr	Phe	Trp	Ile	Pro	Ser	Ala	Gly	Leu	Ser	Glu	Pro	Ala	Leu	734	

139

gcc gac gag gcc ggc gtg gag atg gcg acg ggc ctc gcc ggc ccc atg	830
Ala Asp Glu Ala Gly Val Glu Met Ala Thr Gly Leu Ala Gly Pro Met	
240 245 250	
ccc gcg ttc gtc tat ttc ctc atc gag gcg ttc atc gac gcg ggc gtg	878
Pro Ala Phe Val Tyr Phe Leu Ile Glu Ala Phe Ile Asp Ala Gly Val	
255 260 265	
gcc ctg ggg ctg ccg cgt gat cag gcc acc ctg gcg acg gtg gag agc	926
Ala Leu Gly Leu Pro Arg Asp Gln Ala Thr Leu Ala Thr Val Glu Ser	
270 275 280 285	
atg cgc gga agc ctg gag ttg ctg cat cgc agc cag gag tcc ccg gcg	974
Met Arg Gly Ser Leu Glu Leu Leu His Arg Ser Gln Glu Ser Pro Ala	
290 295 300	
gcg ttg cgc aag cag gtc acg agc cct ggt ggc gcg acg ctg gcc ggc	1022
Ala Leu Arg Lys Gln Val Thr Ser Pro Gly Gly Ala Thr Leu Ala Gly	
305 310 315	
ttg aag gtg ttc gag agc gcg ggc gtc cgg gat gcg atg gac gcg gcg	1070
Leu Lys Val Phe Glu Ser Ala Gly Val Arg Asp Ala Met Asp Ala Ala	
320 325 330	
gtg cgc gcg gtg cac gcg cgc gcg cag gag ctc ggg cgg ctt ccg ctc	1118
Val Arg Ala Val His Ala Arg Ala Gln Glu Leu Gly Arg Leu Pro Leu	
335 340 345	
gcc ccg ctc gag ccc ggg aaa gag ggc tga ctccctggtaa ctccctcgccg	1168
Ala Pro Leu Glu Pro Gly Lys Glu Gly	
350 355	
tacccccctca ttgccgaggt gacggcccgcc tcacgccagc agcgccgaga gcgggaacga	1228
cgcaagcgtg aggcttcgct cgggagcggc ggctaaaccg gaagcggatg agagtccaag	1288
agaaccctc atacgacaac tctcgtagtt ctcccgccc cgaagccttt acgacggagg	1348
aaggacggtg ttcccctacg cgggtccgtg ggcaatccgt gcccgggacc tctggagttc	1408
cccatgcgag gtctgttgca gtccttttcc ctggcgatgg tgttcctcgt gggtctgctc	1468
gcggggccgtc cctcgtcagc ccatgaatcg ctccgtgagc gcatcgatgc cttcgtgcgc	1528
gcgggagcaga aacggatgaa t gtg ccg ggc gtc gcg gtg gcc gtc gtg agc	1579
Val Pro Gly Val Ala Val Ala Val Val Ser	
360 365	
cac ggc agg gtc atc gtg gcg aag ggg tat ggt ttc tcg aac ctg gag	1627
His Gly Arg Val Ile Val Ala Lys Gly Tyr Gly Phe Ser Asp Leu Glu	
370 375 380	
cat cag gtt ccc gtc acc ccg gac acc ctc ttc cag tcc ggc tcc ctg	1675
His Gln Val Pro Val Thr Pro Asp Thr Leu Phe Gln Ser Gly Ser Leu	
385 390 395 400	
ggc aag cag ttc acc gcc atg gcg gtg atg ctc cag gtg gag gcg ggc	1723
Gly Lys Gln Phe Thr Ala Met Ala Val Met Leu Gln Val Glu Ala Gly	
405 410 415	
cgc atc tcg ctg tcc gac agc atc acg aag tac ttc ccc gac gcg ccc	1771
Arg Ile Ser Leu Ser Asp Ser Ile Thr Lys Tyr Phe Pro Asp Ala Pro	
420 425 430	

tcg acc tgg gcc ccc atc acg att cgc cac ctg ctg acg cac acc tcc	1819
Ser Thr Trp Ala Pro Ile Thr Ile Arg His Leu Leu Thr His Thr Ser	
435 440 445	
ggc ctc tcg gat ggg ctg ctc gac ctg cgc aag gac tac acc gac gag	1867
Gly Leu Ser Asp Gly Leu Leu Asp Leu Arg Lys Asp Tyr Thr Asp Glu	
450 455 460	
gag ttc gcg cag ttc ctc tat acg ctg ccc ctg gac ttc ccc gcc ggc	1915
Glu Phe Ala Gln Phe Leu Tyr Thr Leu Pro Leu Asp Phe Pro Ala Gly	
465 470 475 480	
ctg cga tgg aac tac agc aac gcc ggc tac gtc ctg ctg ggc ctc ctg	1963
Leu Arg Trp Asn Tyr Ser Asn Ala Gly Tyr Val Leu Leu Gly Leu Leu	
485 490 495	
gtg aac cgc gtc gcg ggc acc tcc tac gtc aac gtc ctg ggc gag cag	2011
Val Asn Arg Val Ala Gly Thr Ser Tyr Val Asn Val Leu Gly Glu Gln	
500 505 510	
gtc ttc aag ccc acc cgc atg aag acg gcg cga ggc atc agt gag gct	2059
Val Phe Lys Pro Thr Arg Met Lys Thr Ala Arg Gly Ile Ser Glu Ala	
515 520 525	
gac att gtt ccc aat cgc gcg gcc ggt tac cgc ctg ctc gac ggg gtg	2107
Asp Ile Val Pro Asn Arg Ala Ala Gly Tyr Arg Leu Leu Asp Gly Val	
530 535 540	
gtg aag aac cag gag tgg gtc tcc ccc tcc ctc aac acc acg gcg gat	2155
Val Lys Asn Gln Glu Trp Val Ser Pro Ser Leu Asn Thr Thr Ala Asp	
545 550 555 560	
ggc tcg ctc tac ttc tcc ctg aaa gac atg ctg gcg tgg gat gac gcc	2203
Gly Ser Leu Tyr Phe Ser Leu Lys Asp Met Leu Ala Trp Asp Asp Ala	
565 570 575	
gtg gat gac cgc gcc atc ctg acg aac ggg agc tgg cgg gac atc ctg	2251
Val Asp Asp Arg Ala Ile Leu Thr Asn Gly Ser Trp Arg Asp Ile Leu	
580 585 590	
agt ccc gtg aag ctc tcc agc gga gct tcc tat ccg tat ggc ttc gga	2299
Ser Pro Val Lys Leu Ser Ser Gly Ala Ser Tyr Pro Tyr Gly Phe Gly	
595 600 605	
tcg ttc atc gag gag cgc aac ggc cag ctc gtg cac cag cac gcc ggt	2347
Trp Phe Ile Glu Glu Arg Asn Gly Gln Leu Val His Gln His Ala Gly	
610 615 620	
gca tgg cag ggc ttc aag acc gcc tac tac cgc ttc acc gga gac gcc	2395
Ala Trp Gln Gly Phe Lys Thr Ala Tyr Tyr Arg Phe Thr Gly Asp Ala	
625 630 635 640	
ctg tcc atc atc gtg ctg ctc aac ctc gcg gag ccg aac ccc gcg gcc	2443
Leu Ser Ile Ile Val Leu Leu Asn Leu Ala Glu Pro Asn Pro Ala Ala	
645 650 655	
atc gcc aat ggc atc gcc gcc atc gtc aac ccc gcg ctc gcc gtg cct	2491
Ile Ala Asn Gly Ile Ala Ala Ile Val Asn Pro Ala Leu Ala Val Pro	
660 665 670	
ccg ctg gcc ccc att ccg gac ctc gaa ccg gag gtc acg gcc cgg ctc	2539
Pro Leu Ala Pro Ile Pro Asp Leu Glu Pro Glu Val Thr Ala Arg Leu	

675	680	685	
gcg acg ctg ctc gag cag gca cgg gcc gga gcg ctg aac cct gcc gag Ala Thr Leu Leu Glu Gln Ala Arg Ala Gly Ala Leu Asn Pro Ala Glu 690 695 700			2587
ttc gcg tac gtc ccc ggg tgg ttc ttc acg gag gcc gcg ccg tat tac Phe Ala Tyr Val Pro Gly Trp Phe Phe Thr Glu Ala Ala Pro Tyr Tyr 705 710 715 720			2635
cag tcg ctg cta cag agc ctg gga ccg tcg ggg ccg ctc gtc ctg gcg Gln Ser Leu Leu Gln Ser Leu Gly Pro Ser Gly Pro Leu Val Leu Ala 725 730 735			2683
aag cgg gag atg cgg ggc gat gac cgc gtc tac acg tac ctg gtg cag Lys Arg Glu Met Arg Gly Asp Asp Arg Val Tyr Thr Tyr Leu Val Gln 740 745 750			2731
gcc ggc tcc acg atg ctg cgc tac cgg gtg gcg ctc atc ccg gat ggc Ala Gly Ser Thr Met Leu Arg Tyr Arg Val Ala Leu Ile Pro Asp Gly 755 760 765			2779
cgg gtg tcc gcc ttc gcg ctg agc ccg aac taa gagctgaccc gcccgggccg Arg Val Ser Ala Phe Ala Leu Ser Pro Asn 770 775			2832
cacagcggga gcggtggtca gctctggggcc ttcggcagcg tcacggtgcc cggattctcg			2892
atggcgggcga gcgcagggcg ccgcaggtag tggcccgggg cctctcctcg cagccgcgct			2952
gtcttgaggt tcttcctgcc tatcaccatg tcacgcagct ggcgctcgac gtggttattg			3012
tccagcggca cctgcggctg ggtgaggaag accgtcagcc ccgcccagag gttgagcatg			3072
tactcgaggg ccttgcgaaa ggcaactgccg ggcaagtgcgc gctgcgcgtg cgcccagtcg			3132
cggatgcgct cggtcagcgg cgcggaactc tgctgacgca cggcgagccg gtgcgcgagc			3192
gcggcctgac gctcctcgcc ctccagcgcg taccagcccg gcaggtccgc ctcgacggcg			3252
tacagctcgc cgatgaggtc cagcacctct ttgcaggcag gggcaaactg ctgggcttcg			3312
aagaattttc tctcacgtg ggcccagcag aaggccaggg tggctggaca gggcccgtcc			3372
gcgctggact tcgtgcgct ctggtacgcc gcatacccat ccaccagcac gacgccctcg			3432
tagtcgcca gcacggtgcg ccgcgtcgcc ccggagcggc tggggtagat gcgatggaag			3492
acggtgtcgg gcgagggcac cgtccacgcg taccacttgg tgcccggccc cttgtccagc			3552
agcagccaat gcgtctcatc ccgctggtgc aggtgcgcct gggcgctctc gcccttgagc			3612
cgggcattct ccgcgaacgag ccgctcaagg cgctggtgca gccgggcgtt ttccgcctcc			3672
agcagcgccg cccctgctt ggccgtctcc aagtctcgca gctgctcgac gcgcgtcatg			3732
actacctgac aggtgcggca ggacgctcgc tgacgtgcac cggcaccaga gccgcgccct			3792
tctcggcggg cttcggacgg cgggtgtccc cggggcgccc cttgcgccac tcgtagagcg			3852
tcggcccgga cagcccagt tgctgcgccg cgtccgtcac cgtcccaccc ttccgccatg			3912
tgtgctcggc gtagcgcacc gcgaaggcgc gcagcgctc ggggaatggc aacgagccac			3972

cgcgccgccc	ggccttcaat	cgctgcgctt	cctgacggaa	ctgctccagc	tccttctcca	4032
tggaccacct	cctggcgtgt	ggccaagga	agtctcacga	aggctggagc	ccgttacgac	4092
ggggcacggc	gaggtcttac	gactcctgct	cgcgcgccgg	gtgtccgggt	cttccggatg	4152
gagtgggtcc	cctacgccac	gtcagccgac	ctggggccagg	acgtagctga	actcgctctg	4212
agagccgagc	gcggcccgaga	agttgttgaa	gcggccctgg	tccatcgctt	ggcagccggc	4272
ggagtacgtg	tcggtcgttc	cgccgcgatg	gaagtagtag	ccctggccgc	cccagtcgcc	4332
gctgaacaca	tcgtccttcg	agttgatccg	gtgatcgccg	ttgggtgtcg	gccgcacggt	4392
cgggttcccg	tagcgcaggt	ccggcgcgag	gatgttgcca	ccgaagttgc	cgttgaagct	4452
cttccggtag	tggtaggtct	tgttgccgac	gaggcgcccg	accggcttgg	tgctctggct	4512
cgggtcatcg	gccagccggc	ggttgggctc	ggtgttggcc	ttgaactcct	ggacgtgcgg	4572
cttgccgctg	ctgtccttcc	acatgacgac	gagccgatcg	tcgtaggagc	cgccggaggg	4632
gogatcacgg	gtgaagggcg	tgttcaccct	caggccgacg	atgacgcgct	tgccggccgc	4692
gaggtccctg	ttcgcccttg	cgtcgccgta	cttctccacg	tacttcgagt	agagcgcgta	4752
ctccttcgcc	tgagacatgc	cctcggtgtt	cgggatgccc	ttggggagca	cgctggtgtt	4812
tttgttggtc	ggcggcttcg	gtgtcggctt	cggctcaggc	gcgggcgcat	tcacgacctt	4872
cacgtgcagc	ttctggcccc	ggtggatgac	gttgatgttc	tcgacctgcg	ggttcagctt	4932
cttcaatgcg	gccagcgaga	gctcgtgctt	cgccgcgatg	cccagagaacg	tgtcaccgct	4992
cttcacggtg	taatccacgt	acttcttggt	ggccggctgc	tccccggtgc	tgaacctggc	5052
agtgactcca	cccagcgtgg	cgccggaccg	gaccttgga	gtcgtcattc	ggggcactcc	5112
tgaaacccga	ggtgtctggc	tggattgtcg	aaggtgcggg	ctcgaagttt	cttagaattt	5172
gatccatcac	cggccccagt	gccagcaggg	caaacggcga	gtcctttgga	tttctacgct	5232
cagtgaagtg	ctggcggcct	gcggaccggt	cccggggcct	tcacggcggc	aggtagagtg	5292
ggccggctcc	tcatgtgcgt	ccggagcggc	gtcatccgcc	ggagggcagg	gtgatgccga	5352
agtgccgcgt	gagcgcggcc	cgccacgcgg	cctcgtccgg	gagccgctgc	tcgtggcgcg	5412
cgccccccga	ggtgacgacg	agcgcgccct	ccttgagggt	gatgcgcccg	tccggcgtgg	5472
tccgggtgca	gaggcgctgc	tgctgaagt	gggactccgg	cgaggtctgc	tggtggtggc	5532
acatggccgc	gaagtccccg	agctcgcgcg	ggatgagcga	gacgacgtac	tgggccttcc	5592
agtccccga	cgccgcctcg	cgccagacga	cgagccctc	cgcgtcctcg	accagccggt	5652
ggccgcgtcc	gtcctgggtc	tgacgcgcgc	gctcgtcgag	gcgcagcggc	tccaggaagc	5712
actcgccgaa	gccacgctcg	gccagccagc	ggcccgaggc	gtcctccacc	aggagcgcca	5772
gatggctgaa	gtcgggaccg	taggcgctgc	catcgggacg	cgaggccacg	cgcgccgaca	5832

gcagcgtcac	ccggttaaccg	agagtccgga	gcaggcgggc	gaagaggcca	ttgagctcgt	5892
agcagaaccc	accgcgccgc	tggaggacga	ccttctcgaa	gaaggccgcc	tcatcgagcc	5952
ggatggggccg	cttcaggtgg	atgtcgaggt	tctcgaacgg	cacggcctcg	aggtggggccc	6012
gatggagcgc	ggcgagcggc	gtacccgccg	agacaccgat	acgctccagg	tagcgagtgg	6072
cgtcgaacat	gcgcaactgtg	taccgcggtg	ggcccgttgg	gggcgaccca	ttctccaatc	6132
catggggcggc	ccgccattcc	cctggcgcgga	acgcgccggg	gagggctcgc	gaaggaagcc	6192
cggcacgccc	cggagcccga	gcggcgctga	cctacatccc	cgacgcgctc	acgatgagcg	6252
aggacatttt	gtcgttccag	ctgtcgtcga	cgagcgaggc	atcgtccgcc	gtcttggtga	6312
ggctggcacc	gctgaagttg	tcatgttcat	acagggtgac	cttgtagccg	ctggcgacct	6372
tcagggaggt	gatgtcgtca	ttgcgcacgc	cccaggcatg	cagggcgga	aggtgtgaac	6432
ggccaacggg	cagtgccatt	ccatagccac	cgtagtctctg	gtcctggaac	acggtgacag	6492
ccgtggggcc	cgggggggtg	taggtgaagg	ggaagtcacg	ctgagcctgg	gtgaactgcg	6552
cttcccactc	ggcggggccat	ccgaatgcgg	acgtggccag	ggtcttgaga	ttgacgcccg	6612
cggcaccact	ccagaagtgg	atgaactctc	cccagttcag	gcttcgcgcg	tagttgctgc	6672
cgttcttcgg	caggtactgc	gcgagcagga	cgaagtaccg	gttgagcacc	gaggagccgc	6732
cgtggttctt	gtagatggga	tagaaccaat	cccgaacca	gtgggtgttg	gcgcggggga	6792
agctgtccgt	cgtgttgacc	atcaggttgt	accagcgatt	ggcatcactg	gtgcgaccca	6852
acgccaggta	gacgtcgtag	ttgaagatct	ccgcccactt	gctgtcgcgc	cacaggccga	6912
aggcaggcga	gtcgtgcacg	cccttgctgg	cgagttccac	gatgtggggc	acctcatggg	6972
tgacgatgtc	caggtcgttg	ccggtgccgg	cagtccaggc	cgtgggggaa	ctggaaccca	7032
catcaatcac	gttgcgccctg	tcatggctgg	cgtcgaagta	cgtggagggg	tggcctccac	7092
tgtactttcc	tgcgtggaag	atggtgtaga	gctgcgagtc	cgtgccgaag	tgtccgtagg	7152
tcctcttggt	gtatcgccac	acgtcgccca	ggtaggtgtt	gggccagggtg	atggaccggt	7212
tcacggccga	gtcgaagtac	acggccacgt	cattgtcctg	gtagacgcgg	gagacggtct	7272
ggttgtgctc	gaaccagtgc	tccgtccagg	tctcgggcgc	ttgcgcctgg	acggaagtgg	7332
cggagagcat	cggcaacacg	aagcacaacg	aagcaggggg	gaacttcatg	ggggagactc	7392
ctgtgcgcct	ggggatgcgt	gggcttccag	tgacctcccc	atagagaatc	gagccccctt	7452
ttcgggttaa	cgcgaagtgt	atcggaatga	gacggacagc	gagttgacct	ggttcccgtg	7512
ggaacagctc	gcgggcagtt	gagagtgagg	gaaccgcgcg	tcagcgcgtg	gtgcggaagc	7572
ccgcgtaggg	ggacaccgtt	tttttgcgcg	gcgcttcggg	ccagggcatt	cctggggcct	7632
ccctctctct	cgaaggagca	gtaccccg	gtg tcc atc cgt cga acg acc gcc agc			7686

Val Ser Ile Arg Arg Thr Thr Ala Ser
780 785

acc cgc aat gtc acc gtc agt gcg atg ctc gac ccg gcc ctg ggc agc	7734
Thr Arg Asn Val Thr Val Ser Ala Met Leu Asp Pro Ala Leu Gly Ser	
790 795 800	
gcg cgc gac gcg ctc cag gtc ttc ctc gag aaa gcc ctg cgc gcc cag	7782
Ala Arg Asp Ala Leu Gln Val Phe Leu Glu Lys Ala Leu Arg Ala Gln	
805 810 815	
ttc ggc gcc gac tgg cag ggg acg ctg gag cgc gac cgg ctg gcg cag	7830
Phe Gly Ala Asp Trp Gln Gly Thr Leu Glu Arg Asp Arg Leu Ala Gln	
820 825 830 835	
ggc cgt ccc acg agc gag cgc tgg gac tgg ggc agc atc agc tac cgg	7878
Gly Arg Pro Thr Ser Glu Arg Trp Asp Trp Gly Ser Ile Ser Tyr Arg	
840 845 850	
ctc ggg cac ttc cgc tgc ccg cag gcc acg ggc agc ggc gtg gtg acc	7926
Leu Gly His Phe Arg Ser Pro Gln Ala Thr Gly Ser Gly Val Val Thr	
855 860 865	
gag ccc atc gtc gcg ctc gcc cgg ctc gcc aag gcg tat cgc aac ctc	7974
Glu Pro Ile Val Ala Leu Ala Arg Leu Ala Lys Ala Tyr Arg Asn Leu	
870 875 880	
ttc cag cac gag gaa ctc acg gac gcg cag gtg cgc cac gcc atg gag	8022
Phe Gln His Glu Glu Leu Thr Asp Ala Gln Val Arg His Ala Met Glu	
885 890 895	
ggg ctc gcg ctg ctg cag gag gcg ctg ggc aac aag acg cgc gcg gcc	8070
Gly Leu Ala Leu Leu Gln Glu Ala Leu Gly Asn Lys Thr Arg Ala Ala	
900 905 910 915	
gag acg cgc gcc gcg ctc gag gtg tgg gcg ggc ccg ccg gcc cgc gcc	8118
Glu Thr Arg Ala Ala Leu Glu Val Trp Ala Gly Pro Pro Ala Arg Ala	
920 925 930	
gcg aag cgt ccg gcg cgc aag gcc ccg gcg gcc gcg aag cgt gcg aag	8166
Ala Lys Arg Pro Ala Arg Lys Ala Pro Ala Ala Ala Lys Arg Ala Lys	
935 940 945	
ccc gcg gcg aag aag aag gcg ccg aag gcg aag agc gcc cgg cga tag	8214
Pro Ala Ala Lys Lys Lys Ala Pro Lys Ala Lys Ser Ala Arg Arg	
950 955 960	
ggcaggggtgg gccgcggggc ctcagccgtg cgccgaaccc gcgcgggtcg ggccctggag	8274
gggccatcag tggttgtacc gggttgaggg caacgtacgc tcccttgagg tgggtgggcc	8334
gtgggtgagc tacgagcaga cggatctgag aaggcgggtcg tcgcggcatt caaggcggac	8394
agcttctcgt ccgcaggcac atccgtgacg acgctgcgcg aatccacctc gaacaaccac	8454
cgggcctggg tacgaggggtg gggagggcag agtcaggagc cattgggtgc accggggccc	8514
gtgacgtccg tgaacttcgc gccggtgacg ccgcgcgcgt ccagggttc cttgatctcc	8574
tcggccacga tcaaggccac ggtccagccc caggttcgga acaccttggc gccgtccacc	8634
ttcgtcgtgt cgatgcgcat ccgggacacc gaggagtaca cgggcgtgga cgacggggat	8694
gctgaaggcc gcgtgggaat agtcgagcgg tttgccggtg atcttgatgg ggatcctgat	8754

gcgctcgtcc gaaagtgatt tggacgacgc atgcggctcc actcagtgc aagcagaagc 8814
 agccgggttcc cggctacggc ggcctggccg tcacgggccc gtgcggcaac ctgcaacaat 8874
 ccatgagcgt cgaggtgccg cgtatctttc cgggaggcgt cttccctgtg tggaaggggc 8934
 tgctgttcga tcccgcctcg tgggatggct cggacttctt catgccggcg caacgcttcg 8994
 ggttcgtgtt ctgggtggaa gacgatttga gaggatccca ttgcgattga cgtggcacac 9054
 gcgaaatcga tttccctgaa atacttcatt accaggctgt atcacaatca caccagtaaa 9114
 gtaaggctgg cacctccggg gcggcccggg cgcgtggaac ttcacgtctc gaagagaagc 9174
 cagcacacga tgattcgtcc gatcctggcc cggatctgtc acacgaaggg acgcgccggg 9234
 caggagggag ttgggatttc gcgcaactgt ttggcaatgc gcgcgaaaag gtctttcacc 9294
 ggccgcccca cgcttgaggat ggcccagcaa ccatgggatt ggtgggtccg catgccttgg 9354
 caggggggatt gcccggtgga gcatctactt cgctagatgg ttctccacct gacggttggc 9414
 tgctcccaag tctgtgagcg caggtggggc gcacgcacgc aacgcccatt tagatagctg 9474
 cctggcagcc cgcttgcttt ctggctactg cacttccccg atccccctcc atcgtagga 9534
 cgcgcaatgc cttggaggag cggcttgat tttcacggcc ttctacgcac ggagcccaag 9594
 gtcaatcgca cactccaggg caaacgatga atatgaacat gcgagcgcaa gcgggtggaa 9654
 gacacctgag gtgccatgta cagtgcgtga ggaatcccgg ctgtgcatga cactttcgtt 9714
 ggggagggcc gaaacgtgga ctacagggac gcggtacaaa tacaaggaca gacgccagt 9774
 acggcggcgg cagtgcgcga gaagatgcc gtggcagggc tgggcaggta cctgagggcc 9834
 atggagacgc tccgactcgg ggccactcgg gcaagcaggg agttgctgcg cgcctgctg 9894
 gaggtggtgg agcaggattt ggaagacgcg ggtcccatg cacagagcgc ggtgccggtg 9954
 gacaagcatg ggtccccca tccctccggc aagccagtac ctgcctcgcc ctgcaagcgc 10014
 ttccgtgggg tacctccctc tcgtgagggc agtggtgccc ccgagagccc tctgcctgga 10074
 gtgacgaa atg act tcg ggc act cgg aaa aca gga ggc gac ggg gct cgc 10124
 Met Thr Ser Gly Thr Arg Lys Thr Gly Gly Asp Gly Ala Arg
 965 970 975
 cgg aat gac aag tac cgc atc ctc agc ctg gac gga gga gga ctg cgg 10172
 Arg Asn Asp Lys Tyr Arg Ile Leu Ser Leu Asp Gly Gly Gly Leu Arg
 980 985 990
 acg ctg ctg tcc att ggc ctg ttg aag cgc atc gac cag ttg cgg ccc 10220
 Thr Leu Leu Ser Ile Gly Leu Leu Lys Arg Ile Asp Gln Leu Arg Pro
 995 1000 1005
 ggc ttc ctt gat cag gtc gac ctg gtg gcg ggc acc tcg gcg gga 10265
 Gly Phe Leu Asp Gln Val Asp Leu Val Ala Gly Thr Ser Ala Gly
 1010 1015 1020
 gcc atc agc gcg ctc atc atc gcg gcg gct cgc gag cct gcc gtg 10310
 Ala Ile Ser Ala Leu Ile Ile Ala Ala Ala Arg Glu Pro Ala Val
 1025 1030 1035

ggg ctc Gly Leu 1040	gag cag gcg cgg cag Glu Gln Ala Arg Gln 1045	atc tgg ttc acg cct Ile Trp Phe Thr Pro 1050	ggg ctg ttc Gly Leu Phe 10355
gat agc Asp Ser 1055	cca ctc tcc aat cag Pro Leu Ser Asn Gln 1060	ctg ggg gcc ctg gtt Leu Gly Ala Leu Val 1065	ggc cag agc Gly Gln Ser 10400
gcg ctg Ala Leu 1070	atg ccc agc gag aac Met Pro Ser Glu Asn 1075	atg gct cgc gcg ctc Met Ala Arg Ala Leu 1080	acc cac atc Thr His Ile 10445
ctg gga Leu Gly 1085	gac aag acg ctc cgg Asp Lys Thr Leu Arg 1090	gac ctg aag cgc aag Asp Leu Lys Arg Lys 1095	gtg gtg atc Val Val Ile 10490
ccc gcc Pro Ala 1100	ttc cag ctg gac gat Phe Gln Leu Asp Asp 1105	ggc gac cct gat gag Gly Asp Pro Asp Glu 1110	gac agg cgc Asp Arg Arg 10535
ggg tgg Gly Trp 1115	agg ccg cga atc ttc Arg Pro Arg Ile Phe 1120	cac aac ttc ccg ggt His Asn Phe Pro Gly 1125	gac act ttc Asp Thr Phe 10580
gtc aac Val Asn 1130	ctg gat gac tac ctg Leu Asp Asp Tyr Leu 1135	gtg gac ctg gcc ctg Val Asp Leu Ala Leu 1140	cgc tcc agc Arg Ser Ser 10625
tcc ctg Ser Leu 1145	ccg atc gtc tcg ccg Pro Ile Val Ser Pro 1150	gtc cat cag ggc tat Val His Gln Gly Tyr 1155	gtg gat ggg Val Asp Gly 10670
ggc ttg Gly Leu 1160	ttc gcc aac aac ccc Phe Ala Asn Asn Pro 1165	acc atg tca gcc gtg Thr Met Ser Ala Val 1170	gcc cag gcc Ala Gln Ala 10715
atc tac Ile Tyr 1175	gcc aag gcg gcg gat Ala Lys Ala Ala Asp 1180	gtg cgg gac atc ctc Val Arg Asp Ile Leu 1185	gtg ttc tcc Val Phe Ser 10760
ctg ggc Leu Gly 1190	aca ggc gac tcg gtg Thr Gly Asp Ser Val 1195	gac tac ctg gac ggt Asp Tyr Leu Asp Gly 1200	tat aac gag Tyr Asn Glu 10805
aac tgg Asn Trp 1205	ggc tgg aga aag tgg Gly Trp Arg Lys Trp 1210	ctg ctg gac ccg aag Leu Leu Asp Pro Lys 1215	cag ccc atg Gln Pro Met 10850
gcc ttc Ala Phe 1220	gtg gcc gcc acc atc Val Ala Ala Thr Ile 1225	gag gcg gga gtg gag Glu Ala Gly Val Glu 1230	gcc atc gac Ala Ile Asp 10895
ttc cag Phe Gln 1235	gca aag atg ctg ctt Ala Lys Met Leu Leu 1240	ccg cgg ggc aac tac Pro Arg Gly Asn Tyr 1245	tgg agg gag Trp Arg Glu 10940
gat ccg Asp Pro 1250	gtg gtg ccc ttc aac Val Val Pro Phe Asn 1255	ctg ggc aac agt gtg Leu Gly Asn Ser Val 1260	gag gcc cag Glu Ala Gln 10985
atc gcc Ile Ala 11030	acc ttg gat cgg gtc Thr Leu Asp Arg Val 11035	gtc aag agg cac gat Val Lys Arg His Asp 11040	ctg agc aag Leu Ser Lys 11030

1265	1270	1275	
gtg ctg gag tgg gtg gag aac tcg ggc tgg caa ccc aag gcc acg 11075			
Val Leu Glu Trp Val Glu Asn Ser Gly Trp Gln Pro Lys Ala Thr			
1280	1285	1290	
cag gag aca gcc tct ggc gcg cag cgt gag ttc agt ggc gag tcg 11120			
Gln Glu Thr Ala Ser Gly Ala Gln Arg Glu Phe Ser Gly Glu Ser			
1295	1300	1305	
gga gag gcg cgg agc tga gcaggggttc gaacggcctg acggccacct 11168			
Gly Glu Ala Arg Ser			
1310			
gcttgacatg ggggtggccgc agctgcggca tctcgccttc ggaccggccgg aggtggagaa 11228			
atg acc atc atc cat ggg cag cca gga gcc gac ggg gct cgc cgg 11273			
Met Thr Ile Ile His Gly Gln Pro Gly Ala Asp Gly Ala Arg Arg			
1315	1320	1325	
aat ggc aag tac cgc atc ctc agc ctg gac gga gga gga ctg cgg 11318			
Asn Gly Lys Tyr Arg Ile Leu Ser Leu Asp Gly Gly Gly Leu Arg			
1330	1335	1340	
acg ctg ctg tcc att ggc ctg ttg aag cgc atc gac cag ttg cgg 11363			
Thr Leu Leu Ser Ile Gly Leu Leu Lys Arg Ile Asp Gln Leu Arg			
1345	1350	1355	
ccc ggc ttc ctt gat cag gtc gac ctg gtg gcg ggc acc tcg gcg 11408			
Pro Gly Phe Leu Asp Gln Val Asp Leu Val Ala Gly Thr Ser Ala			
1360	1365	1370	
gga gcc atc agc gcg ctc atc atc gcg gcg gct cgc gag cct gcc 11453			
Gly Ala Ile Ser Ala Leu Ile Ile Ala Ala Ala Arg Glu Pro Ala			
1375	1380	1385	
gtg ggg ctc gag cag gcg cgg cag ctc tgg ctc acg cat gag ctg 11498			
Val Gly Leu Glu Gln Ala Arg Gln Leu Trp Leu Thr His Glu Leu			
1390	1395	1400	
ttc cac tca tcg ctc agg acc aga ttg ggg gcc atc att ggc atg 11543			
Phe His Ser Ser Leu Arg Thr Arg Leu Gly Ala Ile Ile Gly Met			
1405	1410	1415	
aac gcg ctg acg ccc agc aag aac atg gct cgt gcg ctc acc cac 11588			
Asn Ala Leu Thr Pro Ser Lys Asn Met Ala Arg Ala Leu Thr His			
1420	1425	1430	
atc ctg gga gac aag acg ctc ggg gac ctg aag cgc aag gtg gtg 11633			
Ile Leu Gly Asp Lys Thr Leu Gly Asp Leu Lys Arg Lys Val Val			
1435	1440	1445	
atc ccc gcc ttc aag ttg gac gat gaa aac cct gac gag aac aga 11678			
Ile Pro Ala Phe Lys Leu Asp Asp Glu Asn Pro Asp Glu Asn Arg			
1450	1455	1460	
cgg gag tgg aag ccg cga gtc ttt cac aac ttt ccg ggt gcc gac 11723			
Arg Glu Trp Lys Pro Arg Val Phe His Asn Phe Pro Gly Ala Asp			
1465	1470	1475	
ttc gtc aac cca gga gac cgc ctg gta gac ctg gcc ctg cgc tcc 11768			
Phe Val Asn Pro Gly Asp Arg Leu Val Asp Leu Ala Leu Arg Ser			
1480	1485	1490	

agc tcc	ctg ccg	atc gtc	tcg ccg	gtg tac	cag ggc	tat gtg	gat	11813
Ser Ser	Leu Pro	Ile Val	Ser Ser	Pro Val	Tyr Gln	Gly Tyr	Val Asp	
1495			1500			1505		
gga ggc	ttg ttc	gcc aac	aac ccc	acc atg	gca gcc	gtg gcc	cag	11858
Gly Gly	Leu Phe	Ala Asn	Asn Pro	Thr Met	Ala Ala	Val Ala	Gln	
1510			1515			1520		
gcg atc	tac gcc	aag gcg	gcg gac	ctg cgg	gac ctc	ctc gtg	ttc	11903
Ala Ile	Tyr Ala	Lys Ala	Ala Asp	Leu Arg	Asp Leu	Leu Val	Phe	
1525			1530			1535		
tcc cag	ggg gcg	ggg gac	atg gtg	cac tac	ctg aag	ggc tac	aac	11948
Ser Gln	Gly Ala	Gly Asp	Met Val	His Tyr	Leu Lys	Gly Tyr	Asn	
1540			1545			1550		
gag aac	tgg ggc	tgg tgg	aag tgg	ctg ttg	aac tcg	aag ttg	ccc	11993
Glu Asn	Trp Gly	Trp Trp	Lys Trp	Leu Leu	Asn Ser	Lys Leu	Pro	
1555			1560			1565		
tgg gcc	ttg gtg	gac atc	gcc gta	gcg gcg	gga gtg	gag gcc	atc	12038
Trp Ala	Leu Val	Asp Ile	Ala Val	Ala Ala	Gly Val	Glu Ala	Ile	
1570			1575			1580		
gac ttc	cag gcg	aag agg	ctg ctg	ccg cga	ggc aat	tat tgg	aga	12083
Asp Phe	Gln Ala	Lys Arg	Leu Leu	Pro Arg	Gly Asn	Tyr Trp	Arg	
1585			1590			1595		
gag gat	ccg cgg	gtg ccc	tct cat	ctg ggc	aac agc	gtg gaa	gcg	12128
Glu Asp	Pro Arg	Val Pro	Ser His	Leu Gly	Asn Ser	Val Glu	Ala	
1600			1605			1610		
cag gtg	gtc acc	ttg gat	cag atc	gtc gaa	aag cac	gat ctg	agc	12173
Gln Val	Val Thr	Leu Asp	Gln Ile	Val Glu	Lys His	Asp Leu	Ser	
1615			1620			1625		
gag gcg	ctg gag	tgg gtg	gac agc	tcg ggc	tgg ctg	cca aag	ggc	12218
Glu Ala	Leu Glu	Trp Val	Asp Ser	Ser Ser	Gly Trp	Leu Pro	Lys Gly	
1630			1635			1640		
atg cat	gag gca	gcc tct	ggc gcg	cag ccc	gaa ctc	agc ggc	gag	12263
Met His	Glu Ala	Ala Ser	Gly Ala	Gln Pro	Glu Leu	Ser Gly	Glu	
1645			1650			1655		
tcg gga	gag gcg	cgg agc	tga gcagggcccg	gagctgctct	gccacccgct			12314
Ser Gly	Glu Ala	Arg Ser						
1660								
ccacgtgagg	tggccgcagc	agcgtcatat	ggtctcccgg	cgccctgcgc	acatccacgc			12374
ccgaggcgct	gagggctccc	cagcccaggt	ctcgcgcgaa	gctccaggcc	tgctcggagg			12434
tgcggaagag	gacgatgogc	tggcggaggc	gacctcgggg	ccgataggag	agcatggctc			12494
gctggtgagc	ccgggtcagc	ctcaacatgg	gggcaatgcc	aggctctcgc	agcagcagca			12554
gccgtgagtc	ctcaggcacc	acccggggcca	tggcggcatg	gtccagcagt	tcctgcacca			12614
gcctctcttt	ccgcaggcgc	tcccatcctc	gggtcaggga	gctgctcagg	cgcgaggcag			12674
cgtctggatt	cagaggagcg	cgagccgcca	ggtagagata	gtcgagcgta	tagggcgcca			12734

ggccgcgcag caccgagccg ctgaagaage gggccgtcgc ccagagactg gctcgctggc 12794
ccgaaggggc agggaaagtcc agcaggctca gcagcgccac ctgctctcca gccgcctcga 12854
gctgtagggc aacctcataa gccaccacgc agccgaagga ccagcctccc aggcggtagg 12914
gccctcgcgg ctgcagctct ctgagctcgc gcacgtacag ggccgccagt tcctccacct 12974
tgtccatggg ctctgcccc tcctgcaacc ctggcggtg gagagcatag aaggggtggg 13034
tgggcgccag gcagcgggcc agcgtgaagt agggaaacac ggtgcccac agcggcgagc 13094
cccagaagaa gggagggctc tcgccctgag gttggatggg caccaggtgg cgccacttcc 13154
gagcgggttc gctggagaca gccggtgacg gctggggagc ggccgctggg attgccccgg 13214
tggccgaggc aggtcgtcc gttgccgggg ctggcatggc ggagagcagt gtgcgcagcg 13274
tagggcgggc gaggaaggag gagagtggca attccacctc cagctccttg cgcaaccgcc 13334
ccatgagctg gatggcgagc agcgagtccc ctttcagctc gaagaagtca tcatccaacc 13394
ccacctcgga gacgccgagc aggtcctgcc agagcgcggc cagggcgcgcc tcgcgctcgg 13454
tgcgaggagc aacatagggc ttggagcgtt cggtcctggc gggggagccg cggggcacccg 13514
agggaaagcg cggctccag gaggagacgc gcaggcgctc gatctccgcc atgcgcgcgc 13574
gcagctcacg ggtggagacg gccacctgcg gcaggccggc cgcgaggatg cgctcgaagg 13634
cctcgagcc ctggcgagg cgagccccga caatcgccac gcctgcccg gggtgcaactg 13694
gccttgagga gctcgggagc gctggcggcc gtgacaaggg ctcttcggg gtgagcacca 13754
ccttgccctac gtgcctgccc ttggagacga gctggagcgc ctggccacg cgctccagt 13814
gccatgccgt gtagggcagc ggctccagct ctccgcgctc gatctgccgc agggcctcgc 13874
ccatcagctc ccggtagccc ttcacccccg ggtctatctg cgccgcgagg aaggtgagcc 13934
cgcgctcgaa gggagccagc ggcagcgtgc cgccagagag catgtccctt acaccagct 13994
cgacgaagcg cccgtggcga gcgagcacgc ccaggctgcg ggagaggaac tcgccgtca 14054
gcgagttgag caccacgtcc acgcctcgcc cctgggtgtg ttgcagaagc tggctcgtcaa 14114
accggggcg gcgcgagtcc atggccagcg ttacccccag ggagcgagc agctcgcgct 14174
tcggctcgct gcccgcggtg gcgtacaccg tggcgccccg ccgcagcgcc aacttcaccc 14234
ccgccagtcc ccccccggtg gccgcggcgt ggatgaggac gtgctctccc gcgcgagcc 14294
gcgccagctc ctccagggca taccaggcgg tgaacagagt gacgggcacc atcgccgcct 14354
gctcgaagct caggcccgct ggcttgggga agacctgctc ctgggggact cgcacatgcg 14414
aggcgaaggc tcctgtgccc gcggccacca ccgcctcgcc cactcgcagg cctctcactc 14474
ctggacctac cgcgagatg atgcccgcgc actcgactcc cagcgccagt ggggcgcgct 14534
tggccagctc cactggcaac gccccagcg ccagcagcgc gtccttgaaa ttgaggccccg 14594
cggcgagcac ctgcacctcc acctctccct cctgcggcgg aaggcgctcc aggggttgca 14654

gtcgcagcga gtcgagcgcg cctggtcgtg gaagcactgc ctggcgcgga gccttgccag 14714
cgggctgtgg ccgagccaac gagtccacct ccaccggcg gagcacgaac tcctccacgc 14774
acagcagctc ccgcccttcg ccgtccatga ggcgcacctc cagccgcacc actccggcct 14834
gggcactggc gggctccagc cggcggtatgt ggctgtacag ctgcccgggc agcggggcgt 14894
ggatgcgggt gttgccatag gacagcgga gccaggcacc ggggggcagc ggtgcaaacc 14954
cggctcgccgc gtccagcagg gcagggtgca ggggccattg ccgcagatcc tcggccagct 15014
cctcggggag ctcgatcaag gccaggccct cttctccctt ccacgcgcgg tgccagctca 15074
gcgaggacca gcgcggccc agggcagggt ccagctcgcg ggggccctga gctggctcat 15134
gctccctcgg cgctgggtc aagccgagcc gttcctccag ctcccgtagc ggggtgggtg 15194
gtcgcgct cccgcgaggc agcatcctca cccggcccat cgcatggctg cgccagcccc 15254
cattcgcgag gccgccagcc cggctgcgga tctccagcgc atagccctca tcctccggtc 15314
gcagggtgaa tcgcacctc acctcctccc ccgcgggcgc ttccagcaga gagagcagca 15374
ccagttgctc cagctccacg ggggcctcgc ccagttcctg gaccagggcg gctgcacca 15434
gctccaggta cgccaccccc ggcagcgtgg ccgctcctc caggcggtgc tcgtccagcc 15494
accagcagcg ctccgctgc catcgacccg tgtactctcc ggggtgccgtg cgctgcagcc 15554
ccagtgctcc cgccgcagc ccagcccca ggcgcattgg ggcgccacc tccagccagc 15614
gatccagtc catggagatg gtgcggcggt ccgtgcgcc tcgctgtgag tgcgcatagc 15674
cgtcgaggaa ggcactggcg gcggcatagt ccgcctgcc cagcgcgccg agcaatgcgc 15734
tcagcgagga gcacaggacg aagaagtcga gtggctcgtc gcgcagcagt tcgtccagca 15794
ccagggtgcc ctgcacctg gaggccagca cgcccgccgc cgactcgcg gtcttgagct 15854
gcatcagccc atctccagc accagcgcg cgtggatcac cccatggagc gcccctggc 15914
ggcgcttcac ctccgagc acctcgcca tgcgctcccg atcggtgacg tccgccacca 15974
ggggcagcac ctgcctccc gcttcttcaa tggccgcag ccggcggtc cgctcctgca 16034
tggccgcgc cccgccact ggctgaagcc cgctctgcc gggcgctgg gcaatgagga 16094
gcacgttgct ccagcgtgag cgctgctcca cggccgctgg cagcacctc acctcggcaa 16154
agcccagct gcgcagggcg tgctcccagt cgctacctc cagcaaggga gagcgggtgc 16214
gcagctcgtc ctcgtagct caccacctt cggccaaacc ccagatcatg tccaccagc 16274
gctgctgctt caccgtctc accaggcaca ggtggccgcc tgggacgagc aggctctcga 16334
ggtgggcccag ggactgagga acgcgcgggg tggcgtgcac cacgttgagg gcaatcacca 16394
gatcatacga gccaggcgga tggccctgct ctgggggagg gcgggagatg tccagcaccg 16454
agggtgtggag gaagggtagg ccctggcggc ggccgactcc ctccatggcg cggacgaagg 16514

agggggccaat atccgtgaag tggtactcca ccgggtggtg ccgcagcagc ggcagcaggg 16574
cctggagcac caccgactg ccgccgcca cctcgaggat gcgcagggta cgcccctgcg 16634
tgcgctcggc atggcgggcg aggaaatgcc gcagcagcgc caccgactgc cccatgctgc 16694
tccactcgac ccgcttcgt gcggagtcct cctcgagcgc gcccgagggg tagagcacac 16754
tgagggcccg caggctccct cggagcacct cagggtagtgc gcgcacacaa tgctcaagca 16814
gctccagcag gcctgacacc tccgggtagc ctctgagcag tcgctcgtgc agtgctcgag 16874
gtgtgggcac gtgggctggc tccacctgac cctcgagcat gccctctccc actcgaatca 16934
gcttcgcctg ctccaggggtg ccgatcatga aggagaagag gcgctcgaag ccgggctgca 16994
gccccaggcg ctacgcaat gcctgcatcg ccctcggtc tccaagtctc agaggcgct 17054
gaagagggaa gagaaagtgg tagaggtagc tggcgcacag caggcgagc gactcttcca 17114
ggccgggact gtcggccagg gagtgcaggg gccgctggtg ctcgatctgc tcggcggcac 17174
cgcgagcca ctcgagctcc tgtgagagcc gcagccgct gggagtggcc.agccatgcgc 17234
tccagctctc gcgcggtggc gtgggcgagc ggcccgtag caccagccgc gccgacagc 17294
gccgcgccag gtgctctgcc agcgcgagcc ctacgcccc caggccgccc acaatcaa 17354
acacccctct ctgccgaagc gggagctgcg aaggctcggg cgggggcagc ggcacgggga 17414
cgaactcgcg caccagcga gccgcgccgc ggtacgccac cgagcgcttc gacgcggtg 17474
actccacctc cgccaccagc cgctccagca gctcctccgc caccgcccct ggcggcaggg 17534
ccatgtccac gaggcggcac gtgcaccgg ggagttcctg gggaagcac gaagcagggc 17594
cctgcaggag cgcttgagc ggctgtatcg gctcgccctc tgccaccgcg tagagctgat 17654
cggtcaccac cgtgagctcc accgcgggcc gggtgccgtg ggctcccagt gcctgggcca 17714
gcgccagcag ggtgtagaag ccgcgctcct gggggctctg gctgctcccc agcagccaga 17774
ggtggagcac cctcgagggc agctctccgg agtccctcag caggcgcttg aagtccctgc 17834
ggcgccccg gggcagggc cagtgcctcg gggtcaggcg ctcgagccgc tcacctggt 17894
cgacgagggt gacggagtgt ccccgggccc ggagccgctc cgcgagcccc'tggctccacc 17954
catgggctc cgccaacaac agcagccggc ccacccgtct cgagggggcc gccgagggca 18014
gagccgcgcg gctccaacct ggcagtaga agtgctggt gacagggctc tcggcaggtg 18074
tcccaggccg ctcgggaggg gaaggctggg gggctgtctc caccagtag cgctggtgct 18134
caaaagggta ggtgggtagc accgcgcgc tccggcgttc gccgcggtg agccagccc 18194
agtccacctc ggctccctcc accacagtc gccgattgc cgtcagcagc cgctccacct 18254
ccgagcttct ctcccggtg gagccagcg tggacaccac ctgcacgggg ccggagcttt 18314
ggcgcgccgc cagccgcccc agtggtggtg ccggccccac ttccaacagc agcccggt 18374
gcacacctc gcacaaccgc tccagcccct gggcaaagcg caccggtgc agcaggtgcc 18434

ggccccagta ctccggtgag gtcgcctcct cggcctctat ccaggtgcc a gtcacattgg 18494
agatgaaggg caggctcggg ggccgcagat gcatccggcg cagccggggcg acgaatgggg 18554
ccaccgcggg ctccatcatc ggcgagtggg aggcgtgaga gacagccaga ggcgtgcacg 18614
cgatgccccg ctgctcgagt cgtgcgcgca gctcggccag gggctcgggtg gggcccgcca 18674
ccacacaggc ctgggggctg ttgtgcgcgg ccagggccag tgccgcgggt aggagcgggt 18734
ccacttcctg ggctgggagc cccacggcga gcatgcccc cgggggcatg ctgccccatga 18794
gctgtccacg tgcggccacc acctccagcg cctcttccag ggagaacacg cccgcgaggc 18854
acgcggccac atattcgccc aggctgtggc ccaacagcgc ggtggggcgt atccccagg 18914
cctccagcag tcggggccagc gcgtactcca cgggtgaagag caccggctgg catacccgcg 18974
tctcccgag gcgctcctgg gccattccc gctcccgcg gccagggtac agcaggcggc 19034
gcacgtccaa cccgagcacg ccctggagtt tcccggcaca gagatccacc tggttcacgga 19094
aggggcgctc ggactcatgg agcgccgctc ccatgcccac gtactgtgag ccctggccccg 19154
gtagaaggaa ggccacggag cgggcctggg ccgcactctg gccacgcagc agcgctcgg 19214
gtcggcgctt ctccagcagc tctaccgcct gggcagggct gctgctcacc agcgctgtgc 19274
gccaggccat gtggcgggcg cctacctgga gcgtgtgggc cacctcctcc agcggtcct 19334
ggggctgctg gcgcaggtgc tcgcgcagcc gggcctcgac ctctccccgg gcgcgctccg 19394
tgcggtgctga cagcggcagg agctggtagg gccgcgctgc ccgcccgggc gccggccggg 19454
gcggcgctc ctccagcacc acatgcacgt tggtgccacc aataccgaag gagctgacgc 19514
ccgcgcggcg cgggccgtcc accggtccc acggctgccc ccgcccgctc acgtagaagg 19574
ggctgcgctc gaagtcgatg cgtgggttg gggaggtgaa gtgcgggcat ggcgggagct 19634
ggcgggcgctc cagcaccagc acggccttga tcaggctggc gatgcccgt gccgcatcca 19694
ggtgtccgat gttggccttg accgagccga gagcgagta gccgcggcgc tccgtgtgct 19754
tgcggaaggc ccgcgtgagg gcctggacct cgatcggatc tcccagggcc gtggcggtgc 19814
cgtgtgcctc cacgtagccg atgctgtccg cggagacggc cgccacctcc agcgctcgg 19874
agatgaccga ggctgcccc tccacgcttg gcgcggtgta gccgatcttg tgtgccccgt 19934
cggtgttcac cgccgagccg cggatgacgg cgtggatgg gtctccgtcc gccagcgct 19994
cctcgagcag cttgagcagc accaccccc cgcggttgcc tcggttggtg cccgcccgcg 20054
ccgcgtcgaa ggggcggcag taccggtccg gtgaggcaat gcctccctcc tggtagagg 20114
agccttgggc ctggggcacg gagagagagg agccgcccgc gagcgccagg gcgcactcac 20174
cattgagcag gctctggcac gcgaggtgga cggccaccag cgagctggag catgccgtct 20234
gcacggccac gctggggccc gcgaggccca gcttgtacga gacgcgggtg ggcaggaagt 20294

ccttgctggtt ggcgatggcg agctggaagt ctcccagcac ctccctcagc cgggcggttc 20354
cgtgcagggg atgcaggtag tagctgttga ggctggcgct gggaagacc cccacgcgcc 20414
cccttgccgc ctcggagcca tggccggccc gctccagcgc ctctgtccgc cactccagga 20474
agagccggtg ctgtgggtcg agcacctccg cctctcgcgg ggagaagcca aagaacgcgg 20534
cgtcgaactg gtccatcctg tccagcaggg cggcgcgccg cacgtaaccg ggttggcgca 20594
ggagcgaggg cccactccg gagggcagga gctgttcgtc cgtgagctgg gtaaaggcgt 20654
ccacgccctc gtggatgatg cgccagtagt cctccagggg gggggcctgg ggggaagcgcc 20714
ccgccatggc cactaccgct acggctccct cgagctgcgt gtcgtggctc atgctgtct 20774
agtcctcgtc ccgttggtgc tggcgggcag tgcgtctctg ctccagcgcc gcgcgcgcga 20834
tcccccgccg gcgtcggcc tgcgtgggag gggaggcctg tgctggcagc tgggcggtc 20894
gaggctccag gtagtccgc agcgagtgc cctgggggta gggaacagg tccatgagcg 20954
tcagctccag gccgagccgc gcgcgcagtg cccgctccac gcgcaccagg aggatggagt 21014
tgccgcccag gtcgaagaag ctggcctcgg gcgcgatgtc ctggtgctcg agcacctccg 21074
cccagacgga ggcgatgagc tgctccagcg ggcgcgaggg gctcgggacc ggagcgggtt 21134
gcggcgccct cagctcacct gcttcgaagc gcgcctcag ctccgagcgg cgcagcttgc 21194
ctccctccgt gcgtgggagc tgggtaggct cgagcgggat gagatgatcg atccgcaccc 21254
ccacctgctc gagcacctgc tggcggtatg gcgacaggag tgaggcaggg ggttgcgag 21314
cgagtcctc cgtgggcacg aagaagaccg ccagctctc cctgctcccc tcgagcatcc 21374
gcaccgccgc cgccgccgc gagggaggcg cactccctc cacctgctcc accgcggact 21434
cgatctcgta gcaggagaag ttggtgccgt ggatgatgac cagatccttg gcgcggcccg 21494
taatggtcag cgctccctcg gagaggaagc ccagatcgcc tgtctccagc cagccgtcgg 21554
cggtagcag cgccgccgag gccttcgggt ccccgtagta gccgggagag atcatctcgc 21614
cgcgacttg aatgcgcccc acgcactcct cgagagcag ctgcgcgtc gcgtccacca 21674
cccgcagggc ggtgccggca atgggggctc ctacatccat cagccgcagc gcggtgggag 21734
aagcgggtgc gcgtgccga agcggaccgc tcaagctcgc gcgtccagg gtgtgcatgc 21794
ggggcgcggt tccgggccgc cgccatagg tcccccca agccgtctcg gtcattcccc 21854
aggcagcgac gaaggcatcc tccgcactc ccgagggagc caacctccgc atcagctcct 21914
ccaccatggg cgcgagatc tgctctccg cgctgagcag cgcgcgacg cctccaggc 21974
tccagcgccg gtccctccac ttgcgcagcc gctccagcag gtggctgtag gcgaagctgg 22034
gcgcccacgt ctgtgcaatg ccataccgct cgagccactc cagccatcgc aggggctcgg 22094
cgagcacctc ctgggtggga gcgtggaagg tgtgcatgcc ggcgcagagc gggcgagggt 22154
ggtactcgga cagggcgcca atgttggtgca gctgcaacca gacgaggccg cgctcgagcg 22214

gctgggactc gtagacgacg ttggtggcct ccaggcgac cagcaggttg cgggtgcgtga 22274
gcatggcgca cttggcettg ccggtggtgc ccgaggtgag gttgagcagg gccacgtcat 22334
ccatgcgccc tgggggcagg ggcgcctcgc caccctggc tctcagctcc tccacggccc 22394
ccaggatgtc ggcggtgggc gagacctgaa gctgcctggc aagatcgggc agcatctcct 22454
gccgggtgag gacccggggc ccacccagca gctgccgggc atgccgcaga cgtgagagcg 22514
cgggactcgc ctccgctctc gcggggcgcg ccggtagcac caggggcgc actcccgcga 22574
gcacgcaggc ccagagtgc tctaccgtct ctctctccct cccagcacc agcaccagac 22634
tgtcgcccgg ccgcaatccc cactgccgca agccggcctg caggcgccgg gcccgctcca 22694
ccagctctgg cagagtccag gccgtgcaag agccctgtgc ctggacgaag ctgaccgtgg 22754
ccgtggggaa gcgctcggcg gtctgtagaa gtgcctcggc cagtgtccga ggggctgtgc 22814
tcgggcgtgg gagggcgga cactccgaca gggcgagaac ctgcttccgt acgggcgggg 22874
ccagcgtctc acccgccagc tccccagtg tccgcgtcgg gtccgccgc accgcttcga 22934
gcaggctcac gtagcgccgc accagttgcg ctacgcctc ctggccatac aggtccacgt 22994
cgtagatgag ggtgccctcg acgcctggg cagactccca gaggtggaac tccagatcga 23054
agcgggccac gcggctgcga agggggaact ccgagagcag caggcctggc agctcggcg 23114
gccgcagcgg gtgtgtgtgc atggcgagga agatctggaa gagcgggtgc cgatccatgc 23174
ggcgccgggg ctggagcacc tccaccaatc gttcgaaagg cacctcctgg tgctcgaggg 23234
cgccgtggag ctgtcggac acctgggcga gccatgaggt gaaggcggtc gagggcgagg 23294
gcttcgcgcg cagcaccagg gtactggcga agcagccgat gaggccttcc agctccttgc 23354
ggttgcggtt ggcgatgggc atgcccagg ccacgtcgtc ctggccggcg caccgggcca 23414
gcagccattg gaaggcggcg agcagcacgg agaagagcga ggtgcgccgc gcttgagcaa 23474
gtgagcgag ggcctgcacc tgggtgtggtg gcagggcgaa ttcgaacgtg gcgccctcct 23534
gcttggacag cgggggcctg ggccgctctg gcggcagggg gctcaggtcc ggcagccctg 23594
ccagctgctc ctgccagtac tcgagcagct cgcgctcgcg cccacccttg agccggctcc 23654
gctcccacgc ggcgaagtcc gcgtactgca gccgcagcgg tggcagcgcc ggtggctccg 23714
cgtgtacagc agccgtgtac agcgcgaca gctcctgcac gaagacgccg agcgaccagc 23774
catccgctgc gatgtggtgc gtggcgaaca ccagcacgtg ttctccggc tccactcgca 23834
ccaggaggaa gcggcacggc ggctctcgt ccagatcaaa ggaggcgagc gctgcgtgct 23894
ccacccggcg gtacacctcg gcctctcgt cctgggcggg cagccgctc aggtcgcgt 23954
gctcgagcgc agccgggtgc ggcgcgagga tgcgctgcag gggttgacct ccggcgagg 24014
ggaacacgt gcgcagcggc tcgtgacgcc ggacgaactc attcaggctg tgcaccaggc 24074

actccacctt cagctttccc gtgagccgga gcgcccgcgtg cacgttgtag aaggggctcg 24134
taggcctgag ctgttgacagg aaccacatgc gctgctgggc ggacgagagc ggcagcccct 24194
cgggcttgcg ctccacgggc atcagcaatg cgcctcgtc caggccctgc tctttcatca 24254
gggcccgcgag tagctcccgc tgctccggcg ccaggctctc cagctcctcg tgcagcttac 24314
tcatgtggtc tccccttget gccgcggcgg gcctcgagca gcgctcgcgc ggcctcgggc 24374
gtgagcgtct tcacctgttg caggagctgc gccagttgcc gcgctccgt tcttccctca 24434
gcgggcaccg aggcgcctc ctgccccga gtggggagta ctccgcccac caggctgtgc 24494
agcagcacgt gctcctctc caagtgaag agccccgcg cctgggcgaa gtccaggctg 24554
ccgatatggc acagccccag cccactggga gccgcgcca gatcgagaag ctgggtcatc 24614
agccctgcct ccagcgtggc gaagtgcagg gcatgttcgg cgtataccgg ctccacggcc 24674
cccatgcggg cgatgaggaa gagggagaac gcggccgagt cgaagatggc ctggttctgg 24734
tgccgatcat ggatgcggcg atccagcccg gctccggcgg acaggagcac cagccgatgg 24794
gtgctgggat cgtaatacca ggtgccgggc tcgagccctc gcgcccggcc gggcttgagg 24854
tggaggtaga gctgcaccgg gtagagcccg cccgcggagg cgtagaggta acgccgcgag 24914
ccctgcacct cccactcgcg caagggggcc agcagccgtc ccaactgctc atgggtcacg 24974
ggctcgaggg agtaggtgcg cacctgcgg cggcgctcga gagctgtcag ctctcgggg 25034
gccggcgcca agccgctgcc tggcagtgcc acctcaggcg ttccatcggg gaagcggcgc 25094
aggcccgggc ggttggcctt gaagcgcccc cactcctcgg cgggcagaag cgtggagcgg 25154
ggcgccgtga gccgcggagc gcctccctct ccttggctct ttgtcttctc ttggtggagc 25214
gtggcgagcc cctgcacggg ggggttctcg tagatgtcgg ccagtgcgg gcggaactgc 25274
agctctgcct ccaggcggtt gatgaggcgg atgagcgcca ccgagtccgc cccaggtcc 25334
agcaggctgt cctggggctc cagcgctcg agcctcagcg cctctttcac cagggtgcc 25394
agccgctcca cgagcgggtc ggccgcgcg gcctgtgcag ccagtccctg tgtctgtgc 25454
ggctcgggga gctgatctcg ggcatcttc ccgttgcggc tgcggggcag ggactccagg 25514
aggacgaagg cggagggcac catgtacgcg ggtagccgct ccgccaggta gcggcgcagc 25574
tcgcctgccg cgggcgtctg acccgagcga ggaaccgcgt aggccaccag ccgccgact 25634
ccccgaggct ctctcgtgc caccaccag ctgcgcgata gcgcccgggtg ctgggcccagc 25694
gccgcttcaa tctccccag ctccacgcg aagccctgca ccttgacctg cagatcctcc 25754
cggccgagga actcgatgct cccttcgggc agcatgcgcc cttgggtcccc ggtgcggtac 25814
agccgctcgc ctgtgcgcgg gtgacggatg aagcgcgttg cggtgagcgg ctcatccctc 25874
cagtactcgc gtgccagccc ctgcgcgcca atgtacagct cggccggcac ccaccagggc 25934
cgtgcctcca gccgcccgtc cagcacgtgg aatcgctgat tggccagcgg catgccatag 25994

ggaatgctct tccactgcgg tgccacctgg ccaatggggt aggcaatgga ccagatggaa 26054
gcctcgggtg gcctcccag gctcaccacc tgcacgtccc tgccgagtgc gcggatgcga 26114
tcggtagcg ccacggggat ccagtcgccg ctgagcatga ccaggcgag cgccgcgggc 26174
agcttcagcc cgcgctgctc ggcgtactcc accagcagca gcatcagcgc cggggtggag 26234
ttccacaccg tgacccgccc agccaccagc cgctcccacc agtgggaggg atccttctcc 26294
gcctccgccc cgggcagcac cagcgctccg cccgcgccga gcagccccag cacgtcgtac 26354
accgacaggt cgaaggtgag cgcggacagg cccagcacc gatcctccgg gccgacgcca 26414
aagcgggtgt tgaggtccac cacggtgttg agcgcggccc ggtgttcaat ggccacaccc 26474
ttgggtttgc cgggtggagc ggaggtatag atgacgtagg cgaggtgctc cgggtgtaccg 26534
cgaggcggtg gcgggggggc ctcggttgca ggctcgagtt cgtccacggc gatcacctgc 26594
acaccggcg gccagggcac ggtgtgcagc aaggaggact gggtgagcac cacgcgggcg 26654
ggccctcct ccaggagctg gtggagccgg agcggcggt gctctggatc caacggcagg 26714
taggccgcgg ccgcctgcag cacgccgagc acggcggtgg cctgctccca gcccttgtgc 26774
atggcaatgg ccaccaactc ctgaggctgc acctccagct cgcgtaggcg cgcggccagc 26834
gcctgggctc gccttgccag ctgcgcgtag ctgagggtgc gctccgggtgc gagcagcgcg 26894
gggagctcgg ggtggagccg cgcctgggtg aagaagccct cctccagccg tccgctgggc 26954
cgcggcgctt ggggtggcgt gttagcgcgc agcaactcac gctgggcagg aggcagcagc 27014
tccggcagct cgcctccca cgcttctcc tctcggcg gtgccccag cagccgctgg 27074
taggcgtgga acatgtcgtc caccatgccc ggagggaaga gcgcctccac cgagtccac 27134
gccaggacga gcgccccctc ctctcgtgg actccgtggt ccagccacac ctggggagtc 27194
tggctgatgc tgtacaccag ttctccctcg aagaaggaga ggctgcctg ggggcgcgc 27254
cgcgcatcca ggctgaggat gctggtgaag acgacgggca tgatggcgcc cggggagcgg 27314
cggccggtgc ggacgagctc gcggatgagc tgcacggcgc tcacgctgcc gtgctccagg 27374
tcccgccata gctgtgcctg gagccgggag gcacgctcgg cgaaggtgct cgccgcgtgt 27434
gcctctacct ccagcaggac caggaggtg aagtcgccc ccagctcgtc cacctgcggg 27494
tgcaggggca agcgtggaa gaggtgagg ttgagggtga agcgcgggtg acggctccag 27554
cgggcaagca cctcggcgaa ggccgcatg caggcgccg agggggtgag gccgtgggcg 27614
cgcgcccgt cctggagccg ggcccagcgg tgtggctoga gccgagcctc gcgggtgggtg 27674
aaccgcgct gctccagctg cgagggttcc ttcaccagg gcagctcggg cggtggcggc 27734
agcgtgtcca gccgagcca ccagtatgc cgggcgcgt catgggcctc gccctcgcgc 27794
agccggcgct ccgccagcac gtagtcgcgg aaggtgatgg ccagcggctc cagcggccgg 27854

cccggctcgt ggtacagctg ggccaagtcc tgccgaagca ccgctgaact ccacgcatcc 27914
agcatcaggg catccatgct catgtggatg cggacgcctc cctcgtaccg gcagagcacc 27974
agctcgaaga gcggccagcg gtcgctgcgc aggacctggt gggccatgcg ctcacgcagc 28034
tggagccggc gggactcggc ctccctggggg gagagcccc gcagctccac cacctccggc 28094
gtgtactccg gcacctgctc gagcacctgc tgccgccccat ccgggagcac caccatgcgc 28154
agcatgtcgt ggcgctgcag gagctggcgc cagcattgga tgaagcgctc cacctccagc 28214
cccgggcttt cgatttcgaa gtagccatgg gcggcgacgc cgccgagctc gaaggccgag 28274
cggcgaccca ccagtaggc ttccctgcaca tccgtgaggg ggaaaggctc gaagcgcgcg 28334
gctggatccg gcacgagcgc tggcgcgggg gcagtgggag ccgcggcctc accctgtcga 28394
gcccgcagag ccgcggccag cgcggcgagg gtgggggtggc tgaagagggg ggccgtgggc 28454
agctccacac ccaggtgcct gcggagatcc gccgccactt ggatgccag cagcgagtct 28514
cctccaagct ggaagaagct gtccgtgggc tgcacgcgct ccacaccag gtgcttgcgc 28574
cacacctgct cgatgagctg ctccagcgag gagggtggca aggcctgcgg ctcgggcgcg 28634
gcggcgcgcc cggtagcag gcgctccttt tcgcgcacgc cctggctgcc gccaggccg 28694
cgcagcagcg agctgaagtc ctgggggggag atgatgagct gcggtgcccc ttgctccaac 28754
gcgcgcagga aggcctccgc accctcttca gaggagatgc cgcccgccag ccgcgggtcc 28814
tcccaaggag gtacggccac cggtagcggg ctggagcccc gagtcatgga cgcgggcaat 28874
gccgtgcgcc cctgcatgga gacggcgacg cgaccgatgt gcagcgcgcg ggccatgaag 28934
ctgaaggctt cggccacctg ccgcgcgggg aagaggcggg tgggcaatgg ctcgagctgg 28994
ccctgggtga gttgcgtggc cacctcctcg agcaccgctc ggaagtccgg gtggtgggga 29054
ccgaagtcaa tggcgggcgaa agtctgccct cgggccaggg ttcgagggcc cacctgctgg 29114
tcgcataca ggtccctctt cccagctcc aggaagcggc cgtgcggggc cagcacggac 29174
aggcccgcca ggagcagctc tcccgccagc gaggtagca ccacatccac ccacgcccc 29234
ccggtgcgct cgcgcacttc gctgacgaag gaggtgctgc gcgagtccag cacgtgagca 29294
atgccgagcg agcgcaggta ctcgcgcttc tgetcaactgc ccgcggtggc caatatctcc 29354
gccccgggtcc gggatgcgag ctgaacggcg gccagacca gccctccggc cgcggcgtgg 29414
atgaggatgc gctcgcccg cgcaaccgg cccacgggtgt gcaggcgcaa gtacgccgtg 29474
gcgaacggca ccatctgggc cggccctcg gcgagcccc gcgaggcggg cctgcgcacc 29534
acctggctct catccaccag cacgtaggag cggaagcagc ctggggccac cgccagcacc 29594
tcgtcccca cgcgagccc gctgacgccc tcgcccacgg cggcaatgcg tccggagcac 29654
tcgcgcccc gtacgctctc ctccgcctcc aatgccggca tcatccccag cgcgcccagc 29714
acgtccagga agttgagccc agcggcctcc acctcgatct ccacctggcg cggccccggg 29774

gcggggcgggg tgcagcggcg caggccaagg ctctccagca gccccggcgt cccacggcg 29834
acttccacgt tgggcgccgc cggctcgaca ctggcgtcgc tgggcttcag cggccgcagc 29894
aggtaacctat cgatccgcac ccgctcccag ccctccagat ccagcagccg cacctgtagc 29954
cggagttgtt gcgcgccagc ctgcgactcc tcgggcaggc tcacgtggct gagcacctgc 30014
ggtggcagcg gaccgtgaat gagcacgcgc tcatagccaa aggcgagcca gggcactccg 30074
ccgggaatgg cgaagctggt ggccatgtcc atgagcgccg ggtggagctc gaagcgctcg 30134
aggtcaccgt ggaagggctc cggtagcgcg agctgggcga gggcctggcg agggccgcgg 30194
cgaacctact ggaagaggcc atgccagcgt gctccggcgt ggacagggcc ttgctcctgg 30254
ggctctccgg gcacgggctg ctcggtgtcg caggcggcgc gcagcagctc gggctggatg 30314
cgctcggcca gccgaggcac ggcgcgcacc cggccccgcg catgctcatt ccatccccgc 30374
gcttctcccg accggctctc caccgcagg gcctgagcct gcccctcggg ctggagcacc 30434
accggacct gcctggattc gcctcgaggt accgtcagag gctccagcag caacagctcc 30494
gccagctcca ccgcctcggc cccagcgcc tgggcacacg ccgcccgcgc cagctccagg 30554
taggccaccc cgggcagtgt tggcacccct tgcagccggt gctcgtccac caccagctcc 30614
tctccacccc tcagcgtgag gccgagccag tgcgtccgc ccggctcgcg ctgttcgcaa 30674
tcgaagagtg ggtgggtgag cggcgtgggc ggcttctcat gggcccccg tcgggtgcgt 30734
gccaccagcc gcatggccgc ccccgatatc ctccacgtgc ccagctccag cggcagcgct 30794
cccgtaagc cctgccggcg cagggtggtgc gcatacgcat cctggaaggc attggccgcg 30854
cagtagtccg cctggccgaa tgagcccagc gccgacgcca gcgaggagca cagggcgaa 30914
caatccaacc cctcgctcgc cagcacctcg tgcagacca tgctgccccg caccttggtg 30974
cgcagggccc gcgaggactg ctcgctgcgtg cggagctgga tgacgccgtc gtcgaacgtg 31034
gcggcgcggt gcagcacgcc gtggatcttc cccagcggg cgtgcacctg ctccaccacc 31094
gagcgcacgg cggcctctc cgccacgtcc gccgtgtaca cctgcacctc ggcgcgcagc 31154
tgctccagct ccagcaggca gcggatgcgg tgctgggtcc tcgggcgctc cgaggggtcg 31214
gcctcgctcc gccacgccgg ctccgggtgc gccggggctt cctccagcca ctgctccac 31274
tgcctcggg gaggggaagg cgagtgggtg agtagcgcca gccgagcacg cggccgctgg 31334
gccagcgctc gcgccagcac caggccgatt ccaccagcc ctctgctac caggtacacg 31394
cctcgctccc gaaggctgcg cagggggcgc gaggcctcga gcctctgccg ggtggcgcg 31454
ctacagctacc gctgcccgtc gcgccaggcc accgggctct ctcccatggg agcggccagc 31514
tgcggggcca ggtcctcac cagcacttcc cgtggctcgc tccggggccg cacgtcaatg 31574
gtggggcact gcacgtgagg gtactccagg ggaagtaccg cacaggcccc cagcagcgtg 31634

gccttgagcg gctcgagcgg ctcccgcggg cccaccgcgt ggacgaagct ggacagcacc 31694
tcgagctgta cagggcgctc agggcgctgg cggccgagcg cacgggccag ccgcagcagg 31754
ctgaagaagc cctgttccag caccgcttcc ccgtccgggg tacaggggcc gggctccgcg 31814
gtgagagccc aggcattggag gatgcgacgg gggagccggc cctgggactg gagcgcctgg 31874
aagagcgctt cccagtcctc ctgctctctg gggcgagggc ggaaggcgtg ctactgagt 31934
tgctcatagt gcccggccgg ctcgacggcg gtgaggagcg cgccggagtc cgcgggccac 31994
agcgcgggcca gcacctgctc ggccaggccc gaggagtcca tgaaggcgag caccgggccc 32054
gccaggggct gtccccccc gctcgtggca ggtgcctgct cccaggtggg cacgtagaac 32114
cagtcttcca gggcctcgcg cgagggctgc tccgggggca caagggacgc gggcgtcggc 32174
tgctgcgcgc gcggctcgcg ggcagcctcc ggttcaatcc agtagcgctg gcgctcgaag 32234
gggtaggggtg gaagggtgtc gcggtggcgt ggcgcggcgt gcagggcgct ccagtccagc 32294
tccaggccgg cctgccacag ctacccacg gcctcttcaa tgcacggcac ggccgtgtcc 32354
ggttgcacgg gaaggctcgc cacctcctgg gtgcggggcg ggcgggtggg gtgccgccgc 32414
gcaagcgcg tgagcgtggt gccggggcct acctccagca gcaactggctc cttgaggctc 32474
cagagcgctt cgagggcctc ggagaaacgt accggctcgc ggagctggcg ggcccagtac 32534
gtaggctccg tggcctcctc gccggtgagc caccggccgg tgaggccgga gacacagggg 32594
aggcgagggg cctgcaggcg catgcgccgc agcagcgctg tgagtggccc ctggcatgcc 32654
tccatggagg cagagtggaa ggcattggaa gtctccaggc gccggctgga cagccccgg 32714
ctctccagcg cggcggtgag cgcctctacc gccggtagag gacccgaggc cagcgaggcc 32774
cgtggcccg taccgcggc gatacacagc tcgctgccca gcagcgccgc cagctcggcc 32834
tccggcagcg gcacggccag catggcccc ggtgggcatc cctgcatgag ctgtccccgg 32894
gctgccacga gctgaagtgc atcttccagg gagaagacac ccgcgaggca cgcgccacc 32954
agctcgccca ggctatggc tgccagcgcc tgcggttgga gccccagcc cagccaggctc 33014
tgcgccaggg catactccag ggtgaagagc gcgggctggg ccaccgcgt ctgcgcgcgc 33074
agccggcgtg cctgctcctc ctgctccgcg gaggggaaga gcacctccct gacatccagc 33134
cccagcgctt gcctcaggag ccgggcacac tgctccaccg cctggcgga ggtgggctcg 33194
gattcatgga gcgcacgggc cgtgccgacc agcggagtgc cctgccctgg aaagagcatg 33254
acgagagggg gccgctctc ctgctgagca agcggctcgg gtggccgcgg ctcgggagcc 33314
tccaggcgct ccatggcctc ggccagggtg cgacagggtga gggcgcgccg gtaggggaag 33374
cgggctcgcc ccacctggag ggtaaaggcc acgtccgcca gttccaggct cgggtggcgg 33434
cgcaggtgtg ccgcgagccg tcctgcgatg tgctcgagcg cgctgggtgt tcgggccgag 33494
agcagcagca cctgggtggg acgggtgggg ccgctggcgg gcggcgggcg agcctcttgc 33554

aagagcgtgt gggcgttggt gccgccaatg ccaaaggagc tcacccccgc cagccgaggc 33614
ccgcgcgggc tctcccacgg gagcggctgg gtgttgacga agaaggggct ctgctccagg 33674
tggagcgcgg gattgggccc ctcgaagtgg agcgtgggag ggatgagccg gtggcgcagg 33734
gcctgcaccg tcttgatcac cccggccact cctgccgccg catccagggtg tccgatgttg 33794
gatttgaccg caccgaggggt acagaaggcg gtgtcagcgg tatgcgcccg gaaggcgcgg 33854
gtgagcgcgg ccacttcgat cgggtctccc aggggagtg cggtgccgtg agcctccacg 33914
taggtgatgt cgcgtgcgct gatgcccgcc gcggcgtgcg ccagggcgat cacctcggat 33974
tggccctcca cgctggggcg ggtgaagcca gcccgggagg cgccgtcgtt gttgagcgcc 34034
cagccgcgga tcaccgcgtg gatgggatcc ctggccttca gcgcgtcctc caggcgcttg 34094
agcaccacca cgcccacgcc atccccctg acgggtgccg tagcacgcgc gtcgaagggg 34154
cggcagtgcc cgtccttcga ggcgactccc ccctcctcgt agaggtagcc ggagcgcgcg 34214
ggcagccgca gcgagacgcc gcccgcgagc gccatgtcac actcgaagcc gcgcagcgcg 34274
ctgcacgcca tggtgatgga gacgagcgag gtggagcacg cgggtgtggat gccgagcgcg 34334
gggcccgcga gattgagctt gtaggccacg cgcgtggcga gaaagtccgg gtagttgcca 34394
aagaagcccc ccacatcctt ggagccgctg agcgggtccc cgtgtccgcg cagcagctgg 34454
aagcggtagc tggtgtcccc ggcgcccacg tagacgcca ccagcctgg gagctgatcc 34514
gggcgaaggg ccgcacctc cagcgctcc cagcagcact ccagcagcag gcgttgctgc 34574
ggatccagggt ggctcgcctc gcgcggcgag tagccgaaga agcccgcgtc gaacagctcc 34634
aggtcttcca gcacaccctt ggcgggcacg tactcggggc gccgacgtat ctgctcggac 34694
acgcccgcct gtctcagctc ttctcggag aagaaggaaa tggactccac tccggccacc 34754
agattgcgc ccagaggact cagctccggg gccccggga agcggccggc catgcctacc 34814
agggcgatgg aagactgcag cgcttcatcc ccactagggtg tcattgggag ctccccgctt 34874
cgaaccggcc cggcggcggtt gggctgcctg ttgctggagc aaggcgcgct ccgtggccat 34934
gcgctgcacc tgcgcctcgg cggctgccag ctctggggg gcctgcagggt gctcggccag 34994
ggcagcaatg gaggagtagc ggaagagggt caccagttcg agccgccggc ccagtcgggc 35054
ctccagccgg ctgcacacct gcgccagcag cagcgagtgc cctcccacgt cgaagaagtt 35114
gtctccccg ctcacttcct ccagtcccag cacttccttc cacacctgcg ccagcaccg 35174
ctccagctcg ttgctggggg gcaccagcgg ggcagattcc agccgacgcg agtcggggac 35234
gggcccggcc gcgcggtcca ccttgccatt gggagtgagc ggcagggact gcagcaccac 35294
gaagcgtgag ggcacatga agtctggtag atgggtctgc aggtggggcc gcagcgcgga 35354
ggccagcctc agtgagaggc tggcccgcg aggtccgag gacagacgcg cggaggctgg 35414

gggcggcgtg cggttgagcg gcgtggggcc gagccaccgc ggggtgcatcc ctggcgctgc 35474
cgggaggaag agcacgtcga agggccccgtc cctccgcccc ggagaccagc tcaactgccgc 35534
cgtgtagccg agcgactcgg ccagttgcca gaaggggtct ggatcctggg cccctgtgcc 35594
gggttgggac aggcgcgggc gcagctcctc caccgcgcgc gcgctgcctg tgctgcccag 35654
cgcctcggcg gcccggaacc cgggcagcac tcgcgcattg ggaatgcctg ccacccccag 35714
tgccctggc gggccctgcc ggagccgagc ccgcagtga tcgaggctca gtccttgctg 35774
ggaccagtcc acccacggca cctcggaggc cggtagcgcc gtgcgcggcc cgatgtggag 35834
cacggcgctg tagcggaagc gtgccatctc attcgggtgc gtgccgcgcg tcagctcgat 35894
atccacgtgg ctaccccggt gatgctgatg ggccagcgcc acgaagagtg ccgggtccac 35954
caccagctcg ttgtcctcca gcaccgcacg ccgcaccgcg tcgcgccacg cctccagcgg 36014
catggagggg gcggcccgtc ccagctccac cgaggcgctg aaggactcca gcagcggcaa 36074
gctccgcaca tccccacga agacgaagcc cccgtcctcc acggcctcca cgcagcgggc 36134
cagcacctgc cgcaggctact cctgcgaggg gaagtactgc accaccgaat tgagcagcac 36194
cgcgtcgaag gagccggccg ccacgccgt cactcgtcc gccgtgcggg gtagcagctc 36254
cacgcggccc ggggtgctccg gggggaggta gcgacgcacg tagtccaacg ccgccgggga 36314
gaaatccgtg cccacgtagc gctcgcagtg cggcagcagg gcaaagagca tcagccccgt 36374
accacagccc acttcgagga tccgtcgtgg cctgagcccg cgcacgcgt ccacccgggtg 36434
gcgcagccag tccctcatct cctccacggc cacgggttgg ccggtgtagc tgctcttcca 36494
gccgacgatg ttgaaggctg gctctccgcc cacgggggct tcgcggtaca tgtgcccgtc 36554
gaacacggac tccactcgt gtaccagccg ctgctcctgc tgcgagccct gctccgagct 36614
ctccaccgcg ggcacggcgt agcagaccag ctcttgccg ccggggccct cgcgccccac 36674
caccaccgcg gactgcaccg ccgggtgggt gagcaccgcg gcctccacct cgcccggtc 36734
gatgcggaag ccgcgtacct tgagctggtg gtccacgcgg ccaggaact ccagcgtccc 36794
atcctccctg cgccgcacca ggtctcccgt ccgatagacg cgagcccccg gcgtccggct 36854
gagcggatcg gggatgaagc gctcggcggg gagctccggg tggtgcaggt acccccaagc 36914
cagcccgggg ccgccacga gcagctcgcc cggggctccg tcgggcacgg gctggagccg 36974
ctcgtccacc acgctcacc agcccaggtc gatgggggtg ccgatgggca cggagccgcc 37034
cagctcctgg gggctgcccc cgcggtgaca gcaggtgaag gtgcagctct cgggtggggcc 37094
gtagccgttg atgagctggc actcggggag ggcctcgcgc agccggcgta cctgcggcat 37154
gggcagcacg tccctccgg cgagcagctg gcgcacgcg cgcaggctct tgagctgctc 37214
ctccaccatg agccggaaga gcccgcggt gagccacagc acggtgacgc cctgggtgct 37274
cagctcctgt ccaggctct cgagcgactg ctgttccgga gggaccagca cgagccttgc 37334

cccgttgagc agcgcccccc agacctcgaa ggtggaggca tcgaactcca gcggtgcgta 37394
gtgcagcatg acctccccgcg gcgagagctg gacgtaggag ggctcatgga tcagccgcag 37454
cacggcacgg tggggaatgc agacgccctt gggctcgccg gtggagcccg aggtgaaaag 37514
cacgtaggcg agcatgtccc cgcccagcga gcgctccggc ccgagcgcgg ggacctgctg 37574
cgcgagggcc ggctcctggg cggcatcgac cagggtcagt ggctgcccc a ggcctgacag 37634
cacctgcca gctgctcct cggcgatgac ggtgcgcacg cccgcggcct gcaccagcag 37694
gctcagccga cgcgctggct ggcgcacatc cagcggtagc taggcggctc ccgccttgag 37754
gatgcctagc aggctgatga tccgctccag cgtgcgctcg gtggccaacc ccactctgtc 37814
gcccgcgcgc acgccagcc aggccaggta gcgggccagg cgttcgagc gctcctccag 37874
ctctgcatag gtgagggtgcc gcccgccttg ggtgacggcc gtggcgccctg gctgccgccg 37934
cgccacctcc gagaagcggg cgccgatgct tgtttcgatc gccgcggcgg gggagggcgg 37994
cggaacctgt cgcagctccc gggcaggcac caggggcagc tccgccacgc ggcgcgccgg 38054
gagctcggcc acgccctcca gcagcgtgca ccaggcgccc accagccgct ccacctccc 38114
cgcgccgtac agcgcgcggt catacacggc gatgcagctc agcccctctt cctcctccca 38174
cagggtgcagc tcgaggtcga agcgcggtgg ggccaccagg tactccagg gctcggcctc 38234
caggccccgc agccgcaacg gctcgcgggg cgcgttctgc agggcgagca gcacctgcac 38294
cagcgggttc tggttggcct ggcgctcggg ctgcaactcc tcgacgagcc tctcgaaaag 38354
cagctcctgg tgggcgtagg cctctagcgt tgtctccctc acccgctca acagctcgag 38414
gaaggagggg ttgccttcca ggcgggtgcg cagcaccagg gtgttgacga agaagccgat 38474
gaggccctcc acctccccac gtgtgcggtt ggcgatgggc gcgccaatgg cgatgtccgt 38534
ctgcccggtg tagcgcgcca gcagcgcatt gaagccggcg agcaacacca tgaagagtgt 38594
ggccccgtgc tcatgcccga gccgcgcgag cgcctgcacc cgtacgccc ggagcgtgaa 38654
gcggtacgtc gcgcggttga agcgctgcac ttccgggcgg gcatgggtccg cgggcaggcg 38714
cagcggggcc aggcgcgcga gccgctccc ccaccaggct agctcctccc gcagcaccgg 38774
cccttgcaac cagccgcgct gccacagcgc gaagtccgcg tactgcacgg gcagcggcgc 38834
caaccccggc tcagcgccat ggagcgccgc ctcgtacagg tgcgacagct ctcggctcca 38894
aacgctcagc gaccagccat cgcccacgat atggtgcagc tcgaggaaga gcagggactg 38954
tgctcttcc aagcgcagca gggacacgcg cagcagcggc cccttgctca gctccagtgg 39014
ccggagcgcc tgctcgcgca tccaccgcaa ggcctcctcc cgcgcggcga tgccatggga 39074
gcgcaggtcc acctgccgca gcgggaccgg agccggcgcg tggatgcgct gccgaggcac 39134
gccgtcgccg ccaggctcga aagtgggtgcg cagcgctca tgccggcgca cgagctgctc 39194

cagggcgcg cccagggcg ccacgtccac ccggccccgc acatgccacg catccaggat 39254
ctggtaggcg gtgctttgtg gcgcgagctg ctggaggaac cacagccgct gctgggcgaa 39314
ggagaggggc agctctggtg gacgtacagc gggcacgagc ggaggggggc ggagctgccg 39374
gtggccgagc aggacctcca cctgggcccgc cagcgccgag agctgagggg cctcgaagag 39434
ggtgcgcagc ggcaactcca cctgcagcgc ctctcgcaag cgggagacaa cccgcgtggc 39494
cagcagggag tgtccgcccc ggtcgaagaa gtcgtcgtgc acggataccc gttgggtgcc 39554
gagcacctcc ctccagatgt ccaccagcag cgcctccacg gggcctcgtg gaggtaccac 39614
cggagcctgc tcgcgggagg cctcctcgaa gctctcgggt agcgccgagc gatccacctt 39674
gccgttgggg gtgaggggca gtgcctccag caggggagaag ctggcgggca gcatgtactc 39734
gggcagtcgg gtgcggagga aggtgcgcag ctctcgggca cctggggcggc caccggcgcg 39794
ggggacgaca taggcgacca gccgcccagc ggccccctcg ccccgagct ccaccaccgc 39854
cgcggaacac tccgagtgt cgcacaggtg cgactcgatc tccgccggct ccacgcggaa 39914
gcctcgcagc ttgacctgcc tgtccacgcg gccagcgaac accacggccc catccgcctg 39974
ccgccgcgc agatctcccg tgcggtagag ccggcctccg ggctccaggc cgtgcggatc 40034
cggcacgaag cgctccgcgg tggcctcggg gtggcgccag tagcccagcg ccagcccatc 40094
acctcccagg tatacctcgc ccatggcgcc caccggcacc ggctgcagcg ctctgtccag 40154
cacgtgcagc cgcgagttgg cgatgggatg acccatgggc acggtatgcg cctgctcggc 40214
caggtgttcc acccgccagg cgggtggagaa ggtggtgttc tccgtggggc cgtacacgtg 40274
tagaaggtgt cctggagcgc cctgtttcag caccgcctc aactggagg catccactgc 40334
ctcgccgcca aagagcaggt agcgaggggt gctgaaggcc tggggttgc tccggggccac 40394
gtggttgaag agggcggtgg tgacgaagag cacggagatc cgctgctcgc gcaggacctc 40454
ggccagccgg cggcctgaa tggcctcctc cgtggcgagc cccaccagcc gtgccccatt 40514
gagcagcgcg cccagacct cgaaggtggc cgctcaaag gaggcattgg aagcctgggc 40574
gacccgatcc tctggcccga gctgaacgta gtccgtgttg cacaccaggc ggggtcatggc 40634
ccggtggcac acggccacgc ccttgggctg gccggtggag ccggacgtat agatgacata 40694
cgccagatct tctccgctgc actggggcag gggctcgcca ccgtcggagg gagcgggggtg 40754
ccatgcctgg agatccagcg tggcgagcc ctggggagcg aaggcgagcg ggccgctgtg 40814
cacgagcaac cgcgcgccgc tgtcctggag catgaactcc tggcgctgta cggggtagga 40874
aggatccagc ggcacgtagg cggccccggc ttggagtacg gccagcatgg actccacctg 40934
ctccggcgag cgctccagcc gaagggcgac caggtctcct cgccgcacgc ctccgggccac 40994
cagacctgt gcgagcccgc gggccccgct cgccaactcg ccgtaggtga ggctgcgctc 41054
gccatacacc actgccacgg ccccgggggc ctggagggca cgctgctgaa agaggccgtg 41114

aacgctggcc tctcgcgat aggggcgctg tgtgcggttc caccgagaga gcacctggtg 41174
gcgctcctgc tctcccagca ccggcagctc cggcacgggc cgatccggcg cctccaccac 41234
cgctccagc acctggcgga agccctccag gagccgcctt gccgtgggct cctcatacag 41294
atccacggcg tactccaacc gcccggtgag cccctccccg gtgtcataca cccacaggtg 41354
gagatccagg tccagcacc cccggggaaa gtcgagcggc tgcgcttgca ccccctccag 41414
cttcagcccg tggccgggca tccatggcgt gtcgctcacc gcgagcgcca cgcgatagag 41474
cgagctgtcc ccgcgctcgc tgccctggctg gagcgccctc accaggtgct cgaagggcaa 41534
ctcctgatgg gcataggcct ccacacacac cctgcgtgcc tgacgcagca actcgcggaa 41594
gcggggggcg cctctcaagt cgagccgcag caggacgatg ttgaggagga agccgaggat 41654
gccctccaag cccggggcgt tgcggttggc ggcggcgctg cccaccgcca tatcgtgctg 41714
gcccgctcac cgttcagca gcacgccag ggcggccagg agcgtcatgt agaggggtggc 41774
gccctcgct tgccccagcg cgcgcagcac ctgggcttgc ggagcggaga gcagggggcg 41834
ccgtacctc gcgcctctca tggaggggat acgcggggcga ggatgatctc ctggtacctc 41894
caggggctg agcccgcgga gctgtcgccg ccagtaagcg agctggacct ggtactctc 41954
tccctccacc caccggcgct gccagtcggc aaggctccgc acctgcaggg gcaggggagg 42014
tagccgcggc tcggccccag cgctcagcgc gccgtagagc gccgccagct cctgttcgag 42074
cacctgcagc ccgactccat ccgcccgat gtggtggagg ctgagcacca gcacgtgctg 42134
ctgctcggcc aggcgcagca ggcgcacgcg caggtggggc tccgtctcca gggcgaagcg 42194
gcgcccggac tcggccaggg ccacctctcg ggcctgctcc agccgctcgc ccacggtggc 42254
accgggtagc gcttccacct ccagcgccgg ctgccacggc tcgccgatga gcgaataggg 42314
gactccgcc tctcgcacca ggggtggtcg cagcgcttcg tgtcgctca ccagcgagc 42374
caccgcccc agcagcgcg ccacgtgcag cgggccggtg aacaccagg ccccgtgag 42434
gttgttgcc ggggcctcag ggtggagctg ctctggaac cacagccgct gctgggcgaa 42494
ggagagcgc agccgccctg ttcgagccac tggggccagg aggtgcctc cctgagagga 42554
gggctggagc tgccgtagca gtgcgatgag ctcatggcga gcgcctcca ggcgggactg 42614
cagctcgggc gtcacaccc ctgggggagc ctggaagcgc agtcgctcgc cctctaccca 42674
caggggtaca cccaggctcg cggcggtggc gagcaacgct ccggcgctca aagttctccc 42734
tctcgatgg aggcagagg aggggtggga gcctgggtgc tccagctctg agcctcgaga 42794
caccgggcca tggccggcag ggtgggctgc tcgaagagca cccgcagggg cagctccagc 42854
caccggctc cgcggatgag cgacaagagc cgtgtggcca gaagcgagt gccccccagc 42914
tggagaagt tgcctccac cccacgcga tgcacgcga gcacgcgccc ccacagctcg 42974

gcaagccggc gctcgggtggg ggtccgcggc gcgaggtggc ctggcgtttg ctcggcttgt 43034
gcgggctcgg ggggaggcag tgccgccacc tccaacttgc cgggtgggctt gagcgggcagc 43094
gcgtccagca gcacgaagga agagggcacc atgtactcgg gcagcttgtg ctcgaggtgg 43154
cttcgcagca ctcccacgag ggtgcgcggc cggctggccg ccaggggctg gttggcgtag 43214
ctcttccagg gtcgaggggg aggcggggga ggcgtcagcc aggcgcggcgg gcgcgaggag 43274
ccagccgcgc gggccatgag cagcgcgtcg aaacggcctt cctcgtcctg ggcgtccag 43334
gtgacgtggc acgtgtagtc ggctgcctcc gccagcctcc agagctgctc tggattccag 43394
gatgtcaccg ccggcagtgc ccgcagccgg ccacacgtgt tgccagtccg cagcagcgcc 43454
agccgtgccg cctcgtccgc cgtgcggggc ttggccacat tgcgcagcac cagccccgcc 43514
ggtcgctccg ccagcatgcg ctccagcctt tccaggccga gcccctcatg cctccagtcc 43574
agctctggca gctcttgcg ggggcctgct ttcgctacgg gggcgagctg cagctccacg 43634
tcgtagcggg agcgggttcat ctcgttcctc ccgctcccgc gcttgagggtg cagccggacc 43694
gcgcccagct ggggaaagtg cgtggcgagc gcggcgaaga agtccgggtc cacgtacaac 43754
cgttcgtcca gcagcatgcg ctggcgggtg cgggcgagca gcgctggcgc ctccaggtcc 43814
gctgacgcct gctccagggc caccgaggca tggaacagct cgaagagccg cagattctgc 43874
acgtcgccaa tgaacagggt tccgcccggg tggagcacgc gcagcgcgcc ctccacgacc 43934
cgcagcaggt agtccacgct ggggaagagc tgcgtcaccg agttgaggat gacgggtgtcg 43994
aagccggctt ccggaagccc ggagaagtca tccgcccggc gggccatgag ccgcacgcgc 44054
tgcgcgagcg gggagccgcc cgtctggagt tgctccctca gccgctcgac ggctaccggg 44114
gagagctccg tgccccagta ggactcgcag cggggggcga gccggcggag caggagcccc 44174
gagccacagc ccagctccag caccgcggc ggaacgagct ccatgagccg ctccactgtc 44234
gtctccaccc actctcgcat ctgctccgga ggcagcggct cgccgcgcac gctgtcattc 44294
caccgcggca ggtcaaaggc cgtgcgcgtt gggccgccat aggtctcatc ccacaccgcg 44354
cgccagcgcct ccacgtgctc tgtctccgcc ccgtccgtgt cctcgagccg tgcggccacg 44414
tagggcacca ggcgcgcctc tccggcgctg tccacgtagg gccgcacgtg cgcttggtgc 44474
acggcggggg gctcgcacag caccacctcg atctcccag gctcgatgcg gtagccccgg 44534
atcttcacct gcgcgtccat ccgttgagg aactccagcg ctccgtcctc tcggaggcgc 44594
gccaggctcc cgggtgcggt gagccgggct ccaggctcgc gggagaaatgg atcgggcacg 44654
aagcgatcag ccgtctgctc cggacgtccc aggtagccgc gcgccagggc ggggcggccc 44714
aggaacacct ctccggccac acccgaggc acgagctgct gctccccgtc cagcaacagc 44774
acccggagggt gggccagtgg ccggcctacg ggcggcagcc gtggccaacg gctcgggtcg 44834
cgtgtcagggt catggctcgt caccacgtgg cactcggagg ggccgtactg gttgtggagg 44894

gtgcactccg gcatccgcga gagcagccgc tgcaactcgg gggtcgcctg gagttgctcc 44954
ccggcggttaa tgagctggcg caggtggcga ggggcccgcgc cgcgctcatc cgccacccgc 45014
gcgagctgct gcagggcgat gaagggcagg tacagccggc tcacctgctc ctgctccagc 45074
acttccagca gccgggcagg atccctgcgc acctcttcgg tcaccatgac cacggtgccc 45134
ccagcggccc aggtggtgaa gagttcctgg tagcagatgt cgaagttgag cgaggcgaac 45194
tggagcgtgc gtgaggggcc acccagccgc tcatgggtga gctgccaggc caccaggttg 45254
gccagacagc gatgcaccat cgcgatgcc ttgggcgtgc ccgtggagcc agagggtgtg 45314
atgacgtagg ccaggttgtc gagccccggc tcgtgtggcg catcctcctg gggctcgcgc 45374
tccagtgcct ccgcctccct ctccagacac aggagcgcgg cgttcgtggg gggcaggcgc 45434
tggagcaggt gctcctgggt gaggaggatg cggcagctcg catctcccag gaggaaggcg 45494
agccgctcgc ggggctgatt gacgtccagc ggaaggaagc cggccccggc cttccatatg 45554
gccagcagtg ccaccggcag ctccaggag cgctccacgc acagccctac cagggtgtcc 45614
gggcccgcgc ccaggcgccg cagccggtgg gcgagcgggt tggcgcgggc gttgagctgc 45674
tcgtaggaga gcacctgcgg gccgaagcac aacgcggggg gatggggggg gagccgcacc 45734
tgggtgctga agcggtggtg ggccagcacc tgcgcgccct ccacgggccc gccaggccg 45794
tgccccgcca gcaaagcctc gtgctcacca ggcgtgagga gcgccagccg cgagagcctc 45854
tcgtgaggcc gctccacgag ctgctccagc aggtgcaggt actgagtggc ggcgcgctgc 45914
atgcgcgcct cgctgaacag gtccgcgctg tagtccagcg cccctccag cccctcaccg 45974
cgcggtcca gcagcagagt gagatcatac gccaccaggc cgggtgaagga ctcccagggc 46034
cggaaggcgc agccgccagc ggtaggctgg ggctgctccg gcagcaggtc gaacagcacc 46094
tggaacagcg ggtggcggtc ggggctgcgg gcaatgtcta gcaggctcac cagccgctcg 46154
aaggggagcg ccgcatgggc gttggcctca cgcacctcct gatcgagctg ccgcagcgcc 46214
tcccgcaccc ggggtgtgagg ctccagcgc gagcgcaatg ccaccggggg gccgaagtat 46274
ccaatgaggc catgcagctc gggccggggg cgctgcgggg cgggccaccc caccaccaat 46334
tcctcctggc gagagtagcg gtgtagcagc gtggtgaagg cggtagcag cagggcaaag 46394
agtgcacgc cctccttggt gccaggctc tccagcgctt ggcgcaggcg cggctgcacg 46454
gtgatggcca cccgccctgc ccggtacgtc tgtctgaggg gacgcgctcc ctccgtgggc 46514
agctccagca ccgtgggcac cccggcgagc cgctgccgcc accaggcctc caaggcctct 46574
ccgcggggcg cgaagcttgc ctcccagcca gcgtagtccc tgtattggac gggcagtggc 46634
gccagctccg ggctcgagcc ttccacgtgc tggccgtaga gcacgccag ctctttgatg 46694
agcacctcgc agctccaggc tccatccgag atcagggtgat gcagggtgac cagcagcaca 46754

tgctccgttt cgcagtcg cgccagccgg actcggaaaca gtgggccttg ctgcagggtg	46814
aagggcgccg cagcctccac gcgcgccagc tccgccaggg cccgagccct gtccgcctgg	46874
gcgacccccg tcagatcgct gagaggcagc tcggggccacc actcggggccg cacacgctgc	46934
tcagggccccg ccgcgccaat ggcaaccacc agccggagcg cctcatggcg cttgacctatg	46994
agccccaacg cccgcgcgcaa cccctctacc tgcaaaggcc cctgtatgtg atgactttgc	47054
aacaaatgct gggccgcggtt gcctggctcg agctgctcca ggaaccacat gcgccgctgt	47114
gaggatgaga ggggagcctt ttcgccgctg gcgtgccgca ccgcgggcag cgcctcgtcc	47174
gacgtggcga gtcccttggc ccggagggct ttcagcgcgt gctcggcagt cgctgtggct	47234
gtcaattgca cccctcctgg aattgaacaa gggggggggc tgtgcccgat agccaatcca	47294
tcgatttcat aaagcccccc acctctattg gcatacctac tgtccagcgt cataccaagt	47354
caagccaccg tccagcttgc tccttctggc atgtcgtga tagtgatgag ttccacagac	47414
gcaacagggtg acccatggag ggggagagcg cc atg gcg gcc cgt tgc cgg cag	47467
Met Ala Ala Arg Cys Arg Gln	
1665 1670	
gca cgc ggg gca tcc atg aca aga att ctt ctt ttc tct ggc aag	47512
Ala Arg Gly Ala Ser Met Thr Arg Ile Leu Leu Phe Ser Gly Lys	
1675 1680 1685	
ggt ggg gta ggc aag acg acg gca tcc gcg gcc acc gcg gtg gcc	47557
Gly Gly Val Gly Lys Thr Thr Ala Ser Ala Ala Thr Ala Val Ala	
1690 1695 1700	
gct gcc aag cgc ggc tac cgc acc ttg gtg atg tcc ttc gac att	47602
Ala Ala Lys Arg Gly Tyr Arg Thr Leu Val Met Ser Phe Asp Ile	
1705 1710 1715	
gcc cac agc ctc tcg gac agc ttc gat ctc gat cgc aag ctc ttc	47647
Ala His Ser Leu Ser Asp Ser Phe Asp Leu Asp Arg Lys Leu Phe	
1720 1725 1730	
gac ttc aac gag gga ctg ccc cag aag gtc gcc ccc aac ctt gaa	47692
Asp Phe Asn Glu Gly Leu Pro Gln Lys Val Ala Pro Asn Leu Glu	
1735 1740 1745	
ctg caa gag atc gac atc cag cac gaa ctc cag cgg cag tgg agt	47737
Leu Gln Glu Ile Asp Ile Gln His Glu Leu Gln Arg Gln Trp Ser	
1750 1755 1760	
gaa gtc tat aac tac atg tcc gtg ctg ctg acc tcc acc ggg gta	47782
Glu Val Tyr Asn Tyr Met Ser Val Leu Leu Thr Ser Thr Gly Val	
1765 1770 1775	
tcc aac atg gtg gcc gag gag gtc gcc att ctc ccg ggc acg gag	47827
Ser Asn Met Val Ala Glu Glu Val Ala Ile Leu Pro Gly Thr Glu	
1780 1785 1790	
gac gtc atc tcg ctc atc tac ctg aac cag tac gtg cag cag ggg	47872
Asp Val Ile Ser Leu Ile Tyr Leu Asn Gln Tyr Val Gln Gln Gly	
1795 1800 1805	
cgc tat gac gtg atc atc gtg gac tgt ccg ccc acg ggc gag tcg	47917

Arg	Tyr	Asp	Val	Ile	Ile	Val	Asp	Cys	Pro	Pro	Thr	Gly	Glu	Ser	
			1810					1815					1820		
ctg	cgc	ttc	gtc	aac	atc	acc	agc	gtg	ctg	cag	tgg	tat	atc	cgc	47962
Leu	Arg	Phe	Val	Asn	Ile	Thr	Ser	Val	Leu	Gln	Trp	Tyr	Ile	Arg	
			1825					1830					1835		
cgg	cgc	ttc	aat	gtg	gat	cgc	acg	ctc	gtc	aag	ctg	gcc	cgg	ccc	48007
Arg	Arg	Phe	Asn	Val	Asp	Arg	Thr	Leu	Val	Lys	Leu	Ala	Arg	Pro	
			1840					1845					1850		
ctg	gcc	agt	cgg	ctc	acg	agc	tat	gat	ctg	ccg	gag	gac	agc	tac	48052
Leu	Ala	Ser	Arg	Leu	Thr	Ser	Tyr	Asp	Leu	Pro	Glu	Asp	Ser	Tyr	
			1855					1860					1865		
ttt	gcc	tcc	ctc	gag	cag	ctc	ttc	gat	cgc	gtc	aag	ggc	att	gaa	48097
Phe	Ala	Ser	Leu	Glu	Gln	Leu	Phe	Asp	Arg	Val	Lys	Gly	Ile	Glu	
			1870					1875					1880		
ggg	ttg	ctc	acg	gat	ggc	aat	cgc	acc	acg	gtg	cga	ttg	gtc	tcc	48142
Gly	Leu	Leu	Thr	Asp	Gly	Asn	Arg	Thr	Thr	Val	Arg	Leu	Val	Ser	
			1885					1890					1895		
agc	gcg	gaa	aaa	atg	gtc	atc	cgc	gag	aca	cag	cgg	gcc	tac	ctc	48187
Ser	Ala	Glu	Lys	Met	Val	Ile	Arg	Glu	Thr	Gln	Arg	Ala	Tyr	Leu	
			1900					1905					1910		
tac	ttc	aac	atg	tat	ggc	atg	acg	gtg	gac	cag	gtc	gtg	gtc	aat	48232
Tyr	Phe	Asn	Met	Tyr	Gly	Met	Thr	Val	Asp	Gln	Val	Val	Val	Asn	
			1915					1920					1925		
cgt	ctg	ctg	ccg	gac	acg	gaa	tac	ttt	gct	cag	tgg	cac	aag	tct	48277
Arg	Leu	Leu	Pro	Asp	Thr	Glu	Tyr	Phe	Ala	Gln	Trp	His	Lys	Ser	
			1930					1935					1940		
caa	tct	gcc	tac	ctg	gat	cac	atc	cag	gag	tac	ttc	tcc	ccc	atg	48322
Gln	Ser	Ala	Tyr	Leu	Asp	His	Ile	Gln	Glu	Tyr	Phe	Ser	Pro	Met	
			1945					1950					1955		
ccg	gtg	gcc	agg	ctc	ccg	ctg	ctg	gag	cac	gag	gtg	gtg	ggg	ctg	48367
Pro	Val	Ala	Arg	Leu	Pro	Leu	Leu	Glu	His	Glu	Val	Val	Gly	Leu	
			1960					1965					1970		
gag	cgg	ttg	gag	gat	ctg	gcg	cag	cgg	ctc	tac	ggg	gac	tcg	gac	48412
Glu	Arg	Leu	Glu	Asp	Leu	Ala	Gln	Arg	Leu	Tyr	Gly	Asp	Ser	Asp	
			1975					1980					1985		
ccc	acc	gag	cgc	tac	atc	tcc	ggg	cag	ccc	tat	cag	ttc	gcc	aag	48457
Pro	Thr	Glu	Arg	Tyr	Ile	Ser	Gly	Gln	Pro	Tyr	Gln	Phe	Ala	Lys	
			1990					1995					2000		
cag	cac	gag	ggc	cgc	tat	cgg	ctg	cag	ctc	ctc	atg	cca	ggt	gcc	48502
Gln	His	Glu	Gly	Arg	Tyr	Arg	Leu	Gln	Leu	Leu	Met	Pro	Gly	Ala	
			2005					2010					2015		
gac	cgc	gga	gac	att	gtc	ctg	gat	cgc	cag	ggc	aat	gac	ctc	atc	48547
Asp	Arg	Gly	Asp	Ile	Val	Leu	Asp	Arg	Gln	Gly	Asn	Asp	Leu	Ile	
			2020					2025					2030		
atc	cgc	gtg	ggc	ggc	ttc	cgc	cgc	cac	gtc	atg	ctg	ccg	cgc	tcg	48592
Ile	Arg	Val	Gly	Gly	Phe	Arg	Arg	His	Val	Met	Leu	Pro	Arg	Ser	
			2035					2040					2045		

gtg gcg cag ctg	gat gcc gtg gac gcg gtc ctg gac aag ggg atg	48637
Val Ala Gln Leu	Asp Ala Val Asp Ala Val Leu Asp Lys Gly Met	
2050	2055 2060	
ttg aag ata gat	ttt gcc cgg ccc tga actggagcc gtg aat gca gtg	48685
Leu Lys Ile Asp	Phe Ala Arg Pro Val Asn Ala Val	
2065	2070	
agg cct gtc atg cag ctg	aaa ggg gcg ggg tca gcc atg cgg atg	48730
Arg Pro Val Met Gln Leu	Lys Gly Ala Gly Ser Ala Met Arg Met	
2075	2080 2085	
aca agg atc ctg ggt gag	cag gaa ctg cgc cag atc gtc cgt gca	48775
Thr Arg Ile Leu Gly Glu	Gln Glu Leu Arg Gln Ile Val Arg Ala	
2090	2095 2100	
acg ggg ctg cac gcg ctg	atg gat gag gtc atc tcc acc ctc gag	48820
Thr Gly Leu His Ala Leu	Met Asp Glu Val Ile Ser Thr Leu Glu	
2105	2110 2115	
tac gag ctg gag cac ttc	gac ccc gcc aac acc gag ctg cgc aag	48865
Tyr Glu Leu Glu His Phe	Asp Pro Ala Asn Thr Glu Leu Arg Lys	
2120	2125 2130	
cgc gag ggc ttc cag tat	cag cac ccc cac ccc ggc gtg ttg gag	48910
Arg Glu Gly Phe Gln Tyr	Gln His Pro His Pro Gly Val Leu Glu	
2135	2140 2145	
tgg atg ccc gtc atg cag	gtg ggg cgc cac gcc gtg atc aag ctg	48955
Trp Met Pro Val Met Gln	Val Gly Arg His Ala Val Ile Lys Leu	
2150	2155 2160	
gtg ggc tac aac ccc cgc	agc ccg gag gag ctg ggc ttg ccc acc	49000
Val Gly Tyr Asn Pro Arg	Ser Pro Glu Glu Leu Gly Leu Pro Thr	
2165	2170 2175	
att ctt tcc tcc ctc agt	ctg tat gat gtg acg acc ggg cac ctg	49045
Ile Leu Ser Ser Leu Ser	Leu Tyr Asp Val Thr Thr Gly His Leu	
2180	2185 2190	
cgg gcc gtt tgt gat ggc	acc ttt cct acc gcg cta cgc acc ggg	49090
Arg Ala Val Cys Asp Gly	Thr Phe Pro Thr Ala Leu Arg Thr Gly	
2195	2200 2205	
gcg gcc tcc gcc gtg gcg	agt cgg ctg ctc gcg cgg gag gac agc	49135
Ala Ala Ser Ala Val Ala	Ser Arg Leu Leu Ala Arg Glu Asp Ser	
2210	2215 2220	
cgg gta ttg ggc ctg gta	ggg tgt ggc gct cag gcc gtc acc cag	49180
Arg Val Leu Gly Leu Val	Gly Cys Gly Ala Gln Ala Val Thr Gln	
2225	2230 2235	
gca cat gcg ctc tcg cgc	gtc ttt cct ctg gga cgg gtg ctg gtg	49225
Ala His Ala Leu Ser Arg	Val Phe Pro Leu Gly Arg Val Leu Val	
2240	2245 2250	
cac gac att gct ccc gaa	gcc atg cga agc ttt gcc cgg cgt gtg	49270
His Asp Ile Ala Pro Glu	Ala Met Arg Ser Phe Ala Arg Arg Val	
2255	2260 2265	
gcc ttc ctg ggc ctg gag	gtg gtg gcg gtc ccc ctg acg cgc gtg	49315
Ala Phe Leu Gly Leu Glu	Val Val Ala Val Pro Leu Thr Arg Val	
2270	2275 2280	

gag Glu 2285	cag Gln 2285	gag Glu 2285	gca Ala 2285	gac Asp 2285	atc Ile 2285	ctc Leu 2290	tgc Cys 2290	acg Thr 2290	gcc Ala 2295	acc Thr 2295	tcc Ser 2295	gtc Val 2295	gcc Ala 2295	gtg Val 2295	49360
ggc Gly 2300	gcg Ala 2300	ggg Gly 2300	ccg Pro 2300	gtg Val 2300	gtg Val 2305	tcc Ser 2305	ggg Gly 2310	cgt Arg 2310	ctg Leu 2310	atc Ile 2310	aag Lys 2310	gcc Ala 2310	cac His 2310	ctc Leu 2310	49405
cat His 2315	atc Ile 2315	aac Asn 2315	gcc Ala 2315	gtg Val 2315	ggc Gly 2320	tca Ser 2320	gat Asp 2320	ctg Leu 2320	ccc Pro 2320	ggc Gly 2325	aag Lys 2325	aca Thr 2325	gag Glu 2325	ctg Leu 2325	49450
ccc Pro 2330	ctg Leu 2330	gag Glu 2330	atg Met 2330	ctg Leu 2335	cgt Arg 2335	cgc Arg 2335	gcc Ala 2340	ctg Leu 2340	gta Val 2340	tgc Cys 2340	ccg Pro 2340	gac Asp 2340	tat Tyr 2340	ctc Leu 2340	49495
ccc Pro 2345	cag Gln 2345	gcc Ala 2345	gtg Val 2345	gtg Val 2350	gag Glu 2350	ggg Gly 2350	gag Glu 2355	tgc Cys 2355	cag Gln 2355	cag Gln 2355	ctc Leu 2355	gag Glu 2355	cct Pro 2355	gga Gly 2355	49540
gag Glu 2360	att Ile 2360	ggc Gly 2360	ccc Pro 2360	agc Ser 2365	ctc Leu 2365	ttt Phe 2365	gaa Glu 2370	ctg Leu 2370	acc Thr 2370	cgc Arg 2370	tct Ser 2370	ccc Pro 2370	gag Glu 2370	cag Gln 2370	49585
tgg Trp 2375	cgc Arg 2375	tct Ser 2375	cac His 2375	cgc Arg 2380	gat Asp 2380	ggc Gly 2380	ctc Leu 2385	acc Thr 2385	gtc Val 2385	ttc Phe 2385	gac Asp 2385	tcc Ser 2385	acc Thr 2385	ggg Gly 2385	49630
ttt Phe 2390	tcc Ser 2390	ttg Leu 2390	gag Glu 2395	gac Asp 2395	aag Lys 2395	gtg Val 2395	gtg Val 2400	ctg Leu 2400	gaa Glu 2400	gtc Val 2400	atc Ile 2400	ctg Leu 2400	gcg Ala 2400	cac His 2400	49675
gcc Ala 2405	gag Glu 2405	cgg Arg 2405	ctg Leu 2410	ggg Gly 2410	gta Val 2410	ggc Gly 2410	cgc Arg 2415	ctg Leu 2415	ctc Leu 2415	cag Gln 2415	gtg Val 2415	gag Glu 2415	tcc Ser 2415	atg Met 2415	49720
gcc Ala 2420	gag Glu 2420	gac Asp 2420	ccc Pro 2425	aag Lys 2425	gat Asp 2425	ccg Pro 2425	tat Tyr 2430	ggc Gly 2430	ttt Phe 2430	ctc Leu 2430	cag Gln 2430	gag Glu 2430	aca Thr 2430	gcg Ala 2430	49765
acc Thr 2435	tgc Cys 2435	ctg Leu 2435	aaa Lys 2440	gct Ala 2440	gcc Ala 2440	cca Pro 2440	gcc Ala 2440	gta Val 2440	taa Val 2440	cgagccaagg	tggttcga	atg Met			49815
aga Arg 2445	ata Ile 2445	ctc Leu 2445	acg Thr 2450	gat Asp 2450	gta Val 2450	ttt Phe 2450	ctg Leu 2455	cgg Arg 2455	aag Lys 2455	tac Tyr 2455	cag Gln 2455	cga Arg 2455	gcc Ala 2455	cct Pro 2455	49860
gga Gly 2460	ccg Pro 2460	cgc Arg 2460	gag Glu 2465	cgt Arg 2465	tcc Ser 2465	ctg Leu 2465	gcc Ala 2470	gcg Ala 2470	gac Asp 2470	gcg Ala 2470	cgc Arg 2470	tcc Ser 2470	aag Lys 2470	ctg Leu 2470	49905
tgg Trp 2475	ctg Leu 2475	tcc Ser 2475	gtt Val 2480	gga Gly 2480	ctg Leu 2480	acc Thr 2480	gcc Ala 2485	ttc Phe 2485	ctc Leu 2485	gtg Val 2485	cct Pro 2485	gcc Ala 2485	tat Tyr 2485	att Ile 2485	49950
gag Glu 2490	cac Tyr 2490	ctc Phe 2495	tgc Cys 2495	cgc Arg 2495	atg Met 2495	ttg Leu 2495	ggg Gly 2500	ctg Leu 2500	ccg Pro 2500	aag Lys 2500	ggc Gly 2500	gag Glu 2500	ctg Leu 2500	ccc Pro 2500	49995
acc Trp 50040	aag Lys 50040	ctc Leu 50040	tat Tyr 50040	gac Asp 50040	atg Met 50040	ctg Leu 50040	cgc Arg 50040	gag Glu 50040	cgc Arg 50040	aat Asn 50040	gag Glu 50040	gcg Ala 50040	cca Pro 50040	gac Asp 50040	

2505	2510	2515	
ctc aag caa cgc ttt cag cag ctc gag cac ctg ccc gag ggt tcc Leu Lys Gln Arg Phe Gln Gln Leu Glu His Leu Pro Glu Gly Ser 2520 2525 2530			50085
ttt ggc cgc gag ttc tgg aag tac tgc cac tcc cat ggc gtg ccg Phe Gly Arg Glu Phe Trp Lys Tyr Cys His Ser His Gly Val Pro 2535 2540 2545			50130
ctg ccg ggt cac cct cac gcc ctg ccc gtg cac tcg gcc gtg cat Leu Pro Gly His Pro His Ala Leu Pro Val His Ser Ala Val His 2550 2555 2560			50175
gac ttc gtc cac ctg ctc tcc ggc tat ggc atc tcc atc tgg gag Asp Phe Val His Leu Leu Ser Gly Tyr Gly Ile Ser Ile Trp Glu 2565 2570 2575			50220
gaa atg ctg acc acg gcc ttc agc ctg ggc ttc atg ttc ggc ccc Glu Met Leu Thr Thr Ala Phe Ser Leu Gly Phe Met Phe Gly Pro 2580 2585 2590			50265
cgc acc cgc aac tat gat ccg gcc cag tac aag tcc ttg gtg ccc Arg Thr Arg Asn Tyr Asp Pro Ala Gln Tyr Lys Ser Leu Val Pro 2595 2600 2605			50310
gtt cgc ctg ctc gcc gag gcg ctt ccc att ccc cag ctg cgg cgt Val Arg Leu Leu Ala Glu Ala Leu Pro Ile Pro Gln Leu Arg Arg 2610 2615 2620			50355
gcg gtg gag cga ggc agc gcg atg agc gtg gac ctc ttc gga gac Ala Val Glu Arg Gly Ser Ala Met Ser Val Asp Leu Phe Gly Asp 2625 2630 2635			50400
tgg gat ccc tgg aag gtc atg gag ctg ccg ctg gaa gag gtc aag Trp Asp Pro Trp Lys Val Met Glu Leu Pro Leu Glu Glu Val Lys 2640 2645 2650			50445
cgg atg tac aac atc cag ccg gag tag ccgctcagcg cgtcgcctcg Arg Met Tyr Asn Ile Gln Pro Glu 2655 2660			50492
gccagcagct cccaggtgcg cgaggcgagc tgctccacct gctcgcgctg catgtgcgcg			50552
gtgttgaaga gaaagctggc ctccaaccgc tgggcaaagg tgatcgccgt cagcaccacg			50612
tcgaactggg ccgcggtcgt ccccgtaag ccaaggccct cgagcgtcag gttcccatag			50672
cgggcggcaa tctccacacg cccacgttg ctgagggcga gcatggagtg caggcccagc			50732
cgtgtcgcgc gttccatcca gggctgggag gagatgtcca cggtagagcg ggatatctcc			50792
aggatcaggc ggaacagcct caagtggctc cgcttcttcc gtgctgcctg cagaggcgat			50852
cggagtctgt gcgccagctc ccagaatggg gtgctgccgc gcgcgcggaa gaagtccgtc			50912
ttccccgtgg tgaacaggcc gaagtcctca cccacgcaa agcgtccccg agcatccaac			50972
gcagaggtga ggggtgagcgt cccccgttc ttggcaccct gcgactcggc ccgagctatc			51032
agcagcgcgg ccccgagcgc tgcgtggacg gaggaccctt gctctcgaca gcgcgcaagc			51092
agccgcgcga gctgctcggg ctccagaatt cgggtggctgt agagcgcctg tccctcggcg			51152

ctgctgcgag ccatgccccg cgggtgcgacg agccgctctg tgagcgtatt gagggtcgga 51212
gacaagcgca gcagcgccga ggagccccgg cgcacggcgg cgccaagcag tcgggtgccc 51272
ggtaccatcg gcccaatcag ctctcatag gcgggccttg gggcgagctc gggcaggtgc 51332
tggtcctggg aggccatgac ctgcagcaga tcgcgcaggg cgaagatgcc ggagtgcgcg 51392
tctccgatca gatgatcata gctgcacagc agctcggagc ggtgctcgcc ccgcacgagg 51452
atgaagcgga cgagcgcccc aggggtccgag ttgaaaggcc tgtcaagctc tgcctcgaca 51512
tgctgttgcc aggactctc tgccttgct tccaccactt ccagcgcaat ggggggcaca 51572
ccttgcgagg tgaagcgtgg cggagtacc gcgacgatct tcgctcggag cagcgggtgt 51632
cgccgttgca gggccaccaa cccctggcgc acagcctgct cctccacgcg tcctcgcagc 51692
ttgcccgtaa tgacgatgtt gaagggggcg ttctgccgta ctttatagaa gacctgctct 51752
gtcgtttccga gggagcgctc catatgatgg atagtaccag agcctgggtg cgcgaagcgc 51812
taccatggag aagcccaaga agagcaagga gaggcatttc ccaggattcc tgaaatgaag 51872
cccacggggc tcaacggatt tcgaaagttg tgatcctcgc ctctgtttcg ttattcggtt 51932
acgaaagggg ggtcgttcgc ttctgacgtg gccctggtct gttccggcca ctctgaagg 51992
ctcagtcctt cgccattatc gacctcatct tccgccgcgt tgtggagcgc atggagcacg 52052
gccgggttca cgtcgcccca tgagaatcga actcaccagg gacgcctctc agcgccctc 52112
accggttttg aagccccgct ccccgatcc gcgaggtgcc cggagggatt gagggtcggc 52172
cggtttccct ctggagcggg gtgggtggag cgtccatgga ggtccacccg cgtccgttca 52232
tcttgggtac agcttgggta cgaggggtgg gacggcaggg gtcaggggcc attggctgca 52292
ccggggacgg tgacgtccgt gaactgggtg ccggagatcc cgcgcgctgt ctcgggcccc 52352
aaagaatgcc taaaaatcg ccttgggtgc agctgcgtcc ttgagaagcc gggtccttat 52412
gccagttttg ttaggcgctc gtatcaccgc tcgtatcacc gcatttaacc actgccttac 52472
gattccaggt cgagcggcag attgagccta accgcgtcct cgacgagcgc ttcgaagtta 52532
gaccgcctgg gctgggcctc caaccacctg tagaactcag ttttcaactg gggatcagtg 52592
gattcgtgta gcagcatctg ttgttcggcc ggattttcca caacgtgttc gaggacagcc 52652
tcgggcggca ggcccagcgc cgtcttcgcc acgtcttgat agctcccttg cgagcccagc 52712
ttcgcattga atgtcttgcg gtcattggcg gtgggtgtctc tgttccatga agcctgcttc 52772
gtggggatct tgatgtgctt cttcgcggcg acgtggacgt gccgctgctg cttctgagct 52832
gggcgctcgg gatccacccg ataccagatg ccagaaatgt gcgagggctc cactcacc 52892
tcattcagga gaagcgccgc atggcggagt tgctcctcag ggatctcgat caagagccaa 52952
ggaaagtcgg cggggacatc gcggtcatca tcggcttcaa acagttcgta tagcgtcatg 53012

ggtcggggttc tccacggggtt ctaaacgaag gctacgcagc ctccgcctca ccttgcaata 53072
tcgttgttgt agacgctgaa tccagcgtcc agccgaggtc aatgctcacc ttgcgaacga 53132
gaccaaacca acactattta gtcgtctggg cgtcggccgg atcgaatgcg accatgagct 53192
tccagtata gttgcctca tccagcgggtg gcagtgggtg gtattcagggc cgaatttgct 53252
tccgcatacca ctccaactgc ggagggccttg ataaatattg ccaaagcccc gaagcagatg 53312
gcgggccctc gttgcgaagg cgcaaactg caatgcttga gatttcggga cccgtggcag 53372
catcaaatac gggaacatga gggctgtcga agtatctaaa caggcaggag tcgtctgccg 53432
tcacgatctc tccagttctt cggcccgtaa agttcatgat gttggtcagt attatggcat 53492
ggatgtttct gcgcgctggc ggagtctcgg ggagtgaac tccctgtgct tggctcctga 53552
actccttcca gtcagtttcg gcaatagaaa cagcccgtag caattggtgc tcggcgcgct 53612
cgatctctc cctcgccccg aactcatcca caggagttgt tgcctcttc aggcacttaa 53672
gctcgaagat aaataggtgc ccttgccaat aaagcaccaa gtcgtattcg atctgccgcc 53732
catctgaagc aggaaatttg atgccccgca cagcgactgc ttgtgttaat tcctcaagcc 53792
tcgattgaat catctgtcga taatactctc cccgggagtc aaacgcggag cttgagcgag 53852
ctgaagctaa atgctctatt gtgcgaatag gatgggaatt ggcaatgatc gaaggaacga 53912
acaacaccct atggctcatct actggtacga gcggctggtc ccagatctct aggttttttc 53972
tctgtgcgtc gaagacgatg ttgttgatat gggaaaagca gaattcgaat tcgatgccccg 54032
acgcctcgct gattaccttg cataggcgct gcttggtgac tgcaaaaaca atgttggtga 54092
ttgctgactt ggtcttgaat gttgcgtgat tcctttgcag cagagccatg gcgagttcat 54152
gcagcgtgaa gttcaacttc aactccttct gtgacttctc cagttcagga tctgctggcg 54212
ctcccattac tgggggttagg aatcgacgcg acatgtagat cttgcgagcc gtggcatatc 54272
gatcaagacc gaggagatct cgcagggtt taatatcaat cttagggtcc catgttttcc 54332
ctggcttcgg gagattgaag ctctcttga ggcgttcag gacttctctt ctactctcca 54392
gttctgcccc aggaccgagt ccatcagaca gcattgcttg ctgcaaaaat aggcgctcgc 54452
gaatgtttcc gatttcgctt tggaggtatt gcgccttggt ttttggtgct cagatgaagc 54512
cgcgtggatc atgctcaggt cgccagtcgg tatactggat ttggttccat agcaggctga 54572
atcgatgcca cgcgtgtgcg caagcgagga gcctgccaaa gtcaggatcg tcataagggg 54632
tctgggggggt aaacttttca tttggggggt ggggtgtcgat cctgtgcatg gagatgcgga 54692
gtattagttc ggcggcttcg gcggtagatt caaaaacaag atcgacactt gtcgttgccg 54752
tcccatcgggt tcctggaacc gcgattgaag ggctcgccat gacaaatggc gagcccgggtg 54812
gggcccgtt gagcttcttt cgaatcgat tgctaacaaa tatgtattgg gattcaacgt 54872
acgtaaatac gcgccacatc attgtgaggt gcgaagattt ccagaattct gatttgctcg 54932

actctttgcga tacgatctgc atggcccttg cgatgctttc gtatccgcgt agcgtgtgct 54992
gctctttctgg cgcaagatct tctgccaatc tgtttattgc ggatactttt tcgaagaatg 55052
ttggattcctt aggatctatg gggcccagtg cgagggcgat caaattcagc cactgagatg 55112
gattctccag tctgatggac tggatgagtg cttgaatctt ttctgagctt ccttcgttta 55172
gggtggctcg cagcttcttg attattttct tcgtttgctc ggctgtagac atgggggcttg 55232
gtttggatgg tgtaaggatg tagttattgg ttttcaggtg tttgaacggc ccatgacacg 55292
cacgggcgcg gacttcatgc tatgtgggtg ctagcttgat agtgaggatg ttctcaacat 55352
tatgggaatg tcaccccgga tagggatgatc gtgcgcagcc cgttcgattc ccacagtttg 55412
agggtgaacg tgtcctgggc cgcgtccggc tcagcctctg cttccaccac gaggcgcgcg 55472
atccgggcat tggcaacggg cgcttcctgc catacagggg gcatgctcag cgtttcacct 55532
gtcttgctca ccaacgatgc tccctgcgcg atccagcgtt gcgcggcttc tgggtagtcc 55592
aaatccatga gcacggacac gcggttggcg gagcggtagc tccatgcacg aagaatcctg 55652
atcgcgtctc caggacgctg gcggacttcc ttctcgaagt cgaggggctt gaccgaaacg 55712
cccgacttgc cgatcacgcg gttcgcaagg aggccagcga ggccgcccgg tcccgcatac 55772
tcggcgcgca ggcgctcgtt ctcttcgtga cacctttggg cctccgctcg ggcttctttc 55832
gattcctgtt gataggactc caccgtgcgc tggttgcgcc acacctcgac caagcgtcc 55892
gcctgagcgg ggtgcgcgac gaggatgaat gccatgccta ccggcgcgtc tccgtccagg 55952
aaccgcaccg tcactcgcag gcgttcgcgc ggcaggactt gctccgacgg caccagccgg 56012
agagtgcctc gtccaatgtc cacgagcgag aagcggtcgg cgccttcaac cgcgaccttc 56072
tcctgaagga gggcggagtt gaactcgaac gtggtggaac cccccgggct tatctgtacc 56132
tccggtgtcg tcattgcctc gaagttctcg gccgagaact caatccggcg taccaccaag 56192
cccgtggctg ctggacgcga ttgcgcggcg gctgccccgg agaaaagcgc tatcaccagg 56252
agaacaacga tgggtgacat gggcatgcgc ggagaccttg aagtagcaga gtggaacct 56312
gcattgatac ttggttcagc cccatgttgc tcctggctcg ggttcactcg aagcggcgga 56372
cggccttcac gtcaaaatag ggatagaccg aggcgctgtt ggagccgtcg ttgatctcct 56432
tcttgaaccc aatagagtcg ttctcaatga tctgcatgca tacagggaag cggtcacct 56492
ttggcgtgcg tgcttcagtt agtcggcgt aaacgcgac cctcctcgatg atgaggcgtc 56552
cggagagcgt ggtccgggccc ggtagatttc ccaatggtct ccccaggatga accgttgctg 56612
ggccccgcgc caccgtgatg ttctgctaa cccctcctc ggcggtcgcg aagaagggtg 56672
ctaccacatc ctgctcatac tcaattcgaa gcctgctcat ggctcgaca gcacctggcg 56732
ggcacgcctc tggcggaggc gtcggggcgga cctggggggc agggcaggcc atcgctaggc 56792

agccccgcgac accgaggccg agcgccttgc caacgggttcc gccccggaag cgctgcggct 56852
cgggctcggga gaatacgccg gtgttcttct tcatcttcac ggaagccttc tcctcggaga 56912
tcgtcgcggg gagggcgacg gccgctaggg tgggcgcgct atcgagcgcg gctgcggctt 56972
cctgagcttc cgggtgccatc catggcgccg ccacttctcg gccggggggcc cagggcgcgga 57032
ggtcgaactc ggggggatgca ggtggtgagt ccacggcatc actgggggtga ggctccaggg 57092
cgtgacgtgc ggcaggagca ggggaaggcg ccaagggtgcg cgcgatcacc agcgcgcccc 57152
cgctcaagac caccagcaga cccatccagg tcgccacacg cagggccttg aagtgaccgg 57212
cttgtgccgg gggcggcgag ggcattggggc cggggcgagc aggaactggg ggtggcagct 57272
ctgcctggat ggccaccgcg tttgagtccg gctcgtgtgc gccctcgccg tgtggcggac 57332
gtggcccccg acgaggccgg gccttgcgct tcttcagcca ttgcaccagc tcgtcttcgc 57392
ccgcgtgcgg caccgccgcg agcgtggtgg cgggtgtccgg gccataggcg tcgcagagtt 57452
tcacgtccca ggactcgtcc gcctggggcca gcagggcctc cagctccgcg cacagcgcg 57512
cggcgtccgg gtggcgctcc tcgggctcct tggccagcag gcgcatgcac acggcactca 57572
gggcctcggg gacgcgtgga ttgagcacat ggggtggttt gggggcgcgg ttgagcacgg 57632
cctccacccc tcgggcctcg tcggggagga aggggtgcct gcccgtcagg agccaataga 57692
gaacgacgcc cagcgagtac aggtcatccg agggggccgg ctggtaggag tggccacgct 57752
cgtccccgtg ctctgctgg aagcgccagg cctcggggcg gcggtactcc ggcggtcccc 57812
gtggcagcac gccgccggtg atgctggggg cgctctcgta cccgcccgcg ccgaagtcca 57872
ccaccactgc ctctccgtcg gaggtgcgct cgatgacgtt gctctccttc aggtctcgat 57932
gcaccacctt cgcccggtgc gcggcgccca gcccgcgcg cagccacgc accttgagca 57992
cgacttctcg ggccgagggg ttctcctctc tggccatac gtccaaccgg cgcccttcca 58052
cgtagtccat gacgatgaag aaggattgag gcgcctcatc cggccactgt ccatgcccac 58112
ggatgcgcac cagattggag tgcttgagcc ggaggaggat ggccacttcg cgctcggccc 58172
actcagccag cccccagagt gggatgagct tgagcgcgta gagccgccct ccacgccggg 58232
cgcggtacac ggcaccgaag cccccggctc ccagcttctc ctccaccagg taccgccga 58292
tgtccgtgcc cgggggggacg acgggatgca gacgcgggct caagtgtcct ctcttcggc 58352
ggggcggtgtg gggtcggcgt tcctggacat gctctcaacc ggtagcggat ctcgtagag 58412
ttccccccgc gaagtggcg gcggagaaggg taaaccttga gcggatacta tcgttgaacc 58472
tggtgagctg tatgctcgga gtgccgcca cggggcattc tcctctttct ccaggtctc 58532
ctcatgtacg cgaagctctc ccagcagatc ggaaccaatg tccgcgccgc ccggcatcgg 58592
gcgggcctgt cgcaggccca ggtggccgag gccatccacg tgccgacgct ggtgttcagc 58652
cgccctggagc gtggccgggt gctgcccagc ctccccacgc tgggtgggcct gtgcgggtgc 58712

ctccgtgtcc ccgtggactt catcctgggc ggcctggccc tggaggttcc cgccccgat 58772
gggtccggcc aataggggcg ccgcgtgacg gggccctgct cgtggatcga catggccgcc 58832
gagatgcggc aggaaggggc gtccccgagg cttgggcctt cctgcccgcg ccaaacctgg 58892
ttctcgggtgt cctggtaccc cgacgcata gaagagct ggcgaccgt attggaagtg 58952
ccgcccgtga gggccggacg cagctcgggc tcacgcaggc ggaggtggcc gagaagctgg 59012
gcctcgcgca catggtctac agccgcctgg agcgcgggaa gatgctgccc agcgtccaga 59072
cgctgctgcg gatgtgctcg gtgctgcgca tctcctcgga cgagttgctg gggatcgcgg 59132
acgcggagga ggggacgaag tcgggcaagg gggagcgagg gcagggcggg gcgcccgaagc 59192
tgcccgagct caccggactg gcgcgcaaga tggacgagga ccagctcgat gccctcgtga 59252
agggtggcca ggtgttgttg cgctgaaacg acgcgggccc acaccggggc ccgcttgctt 59312
caaggaatgc accgccgccc gttctcactc ttcgttcggg gcgggcgtcg gctcgtcggc 59372
gtccaggccc tcggggaaga tgccttcttc gtccagttcc acccggtcgc gcaggacgga 59432
cgctccgac agcatccgcg ccagctctc gctcgcctcc ttccggccga agggctgtga 59492
ctccgccgcg agcgcgcgtc gcagcgcata gcgcatactg gcggccgtgg cgtagcgtc 59552
ggagagcttt cgctggaggg ccttgtggac gatggccttg agcgcgtccg gcaggcccgc 59612
catcgcgttc tccacgtcct cggggcggtg gcggttgacg agcgcgatca tctgcgtgag 59672
gggcaccgag ggcgtctcct ccgtcttcac gtccacgggtg gcgttggggg ccttttcgtc 59732
ctccacgtcg aagaggtgct tgcacgtcag catctccatg aggacgaggc ccagcgagaa 59792
gatgtcggac cggccatcca tgggcttacg gttcaggtag tcgggcgagg cgtacgccac 59852
gtcgccttg agcaggaggc cgggcgtctc ctccctgccc accatgagcg agtaggtagc 59912
gccgaagtcc gtcaccttca cctcgccgct gcgggcccac cgatgttcc ggggcgccac 59972
gtcgcgggtg atgatgccc gcggccggtt ctccgaatct ctacgcgtgt gcgcgtggtg 60032
gagggcgtcg gccacctcgg cggcgatgta gagggcgaag ggcgcgaga gcggcttctc 60092
gcgcaggtg gcgaggctga tgatggtgtc cagcgtgggg ccgtccacgt actcggcgat 60152
gacgtgcggc cggtcggcat ggattttcag gtgatggacc tgggcgatgg ccgggtggtg 60212
gagggcgaag gagagctgga cctcctcgat cagccgctgg cagcgtcga aagtgcggg 60272
attgcgcagg cgcttgatgg tgacgagccc ggcgagtcca tgggggaggt gacgctcggc 60332
caggagcacc acctcgccgt taccgcgcg ctccagcgtg cggagggtct cgtacgtggt 60392
gggcctcgatg gtgaaccgga gtcggggggc ggcgccttc aggtcggagg tgttctcgg 60452
gggcggggc tcgatctgcg ttctcgtggt cggcgtgctg tcatcggtca tgtgcggcct 60512
gcacgtcaga agggacgcgc ggcgggcact ggctcggggc ttccggctcc ctggtggggc 60572

ctcggctgag ggggacgggg cccgtggctc tcgcggtccg aggtctacta ttcataaacc 60632
 caaaggggtat tacaagcgac cacagggtc ggcagggtga gcgcggcgga gcgcctgcct 60692
 ggtgctcgcg ctctgcctct gctccgtgcc ctgagtccca cctctaagca ctggccgcag 60752
 gagcttggcc ttcgcacgtt gaagggtgctg gcggaccacc tgggagtgtt gccgctggat 60812
 ctcttgacgt tccccgagga tgacgagcgg caacgcctcg tggaccgcat ccgcacgatg 60872
 tcctcggccg agatcaagaa gctcttgaag gcgctcgcg accaagacca gcggtagatg 60932
 gcgtccgaca tcggccagca ggcgcgaggc catcaatccc aggagaacgc cggctctcttc 60992
 gggggcaccg gctcgggtgat tccccctacg agcacacgac catcgtgatc accggggggag 61052
 ggtccggcaa cctgggtcctg cggcaccacg cgcggtgta ccggctccaa tgctggcggg 61112
 ccggtcgcgg tcggcgaagg ggcttcgggt ccgtttgcca catggcgatc tccaatgccg 61172
 tggccggttg ggtcgctgac ggttgccgcc cggcaaccac cagggaccaa cgcgctcctt 61232
 tcgagcatcg accggagcgt cgcgttttca tcccgaagcc gctgaagctc tgcgtggatc 61292
 tccttctctt ttcgctggcg gtgcagccgt agggtagcgc caggcagccc tggcgcggcc 61352
 gtggccgtga actgctccga ggtcagccc gacctctccc aggccgcgag catcgcgctc 61412
 aggtactccg tgctccagcg ggtgcctcgg ctcatcacg gaagggtgcc gtccggtgta 61472
 cgtcgcgaat cccctggccg tccacggcgc aggatcgcg agcggcgttg atgcccgtag 61532
 ccctcggggc gcaccgtccg aggtgcgcag acacattcga tatattcctt gtagccacac 61592
 cgaacccggc tccgctccag ggcgccgcaa cgacgtagcc agaatcgata caagttcaag 61652
 agccaccctg catttcaagg cgtcgaagca cggttcggc gttgatggga gaggtgtacc 61712
 ca atg gga cat gag gac tcg gac aag gca ctg ccc gac tcg aag gcg 61759
 Met Gly His Glu Asp Ser Asp Lys Ala Leu Pro Asp Ser Lys Ala
 2665 2670 2675
 aac gag gtt gaa tcc agc cgc aac gag gct gag tgg aac gag ctt 61804
 Asn Glu Val Glu Ser Ser Arg Asn Glu Ala Glu Trp Asn Glu Leu
 2680 2685 2690
 ccg atc tcg gcg cga gac tgc gag gac ctg aag aag tcc tgc ctc 61849
 Pro Ile Ser Ala Arg Asp Cys Glu Asp Leu Lys Lys Ser Cys Leu
 2695 2700 2705
 ctg ccc tcg acg gtg cgg aaa ctc ggt tgg cga acg gac acg gac 61894
 Leu Pro Ser Thr Val Arg Lys Leu Gly Trp Arg Thr Asp Thr Asp
 2710 2715 2720
 gag aac aac cgg aag ttg ttg aac tgg ccc gag gaa tgg aag ggg 61939
 Glu Asn Asn Arg Lys Leu Leu Asn Trp Pro Glu Glu Trp Lys Gly
 2725 2730 2735
 gac gct gga ggg ctg gtg ctc gtg tac ccg ggg ctc aac tac gcc 61984
 Asp Ala Gly Gly Leu Val Leu Val Tyr Pro Gly Leu Asn Tyr Ala
 2740 2745 2750
 aag gtc aag ccg cat cgc ccg cgc gtg ctc cgg ccg aag gag gcg 62029

Lys	Val	Lys	Pro	His	Arg	Pro	Arg	Val	Leu	Arg	Pro	Lys	Glu	Ala	
			2755					2760					2765		
cag	ctc	acc	ctc	cca	acc	gcg	acc	ttc	acg	gac	cac	atc	ctg	cgc	62074
Gln	Leu	Thr	Leu	Pro	Thr	Ala	Thr	Phe	Thr	Asp	His	Ile	Leu	Arg	
			2770					2775					2780		
ccg	tac	acc	cag	gat	gag	gtc	gaa	ggg	gag	cca	gcg	aag	gtg	aag	62119
Pro	Tyr	Thr	Gln	Asp	Glu	Val	Glu	Gly	Glu	Pro	Ala	Lys	Val	Lys	
			2785					2790					2795		
aag	tac	gag	gct	cct	ttc	ggg	tac	gcg	ccc	cag	atc	ttc	gac	acc	62164
Lys	Tyr	Glu	Ala	Pro	Phe	Gly	Tyr	Ala	Pro	Gln	Ile	Phe	Asp	Thr	
			2800					2805					2810		
gcg	atc	ctc	tgc	ccg	atg	ctt	caa	gac	gaa	gaa	gtc	ccc	atg	gcg	62209
Ala	Ile	Leu	Cys	Pro	Met	Leu	Gln	Asp	Glu	Glu	Val	Pro	Met	Ala	
			2815					2820					2825		
ctc	att	gag	ggc	gag	aag	aag	gcg	gcg	ctg	ctc	gcg	cag	gaa	ggc	62254
Leu	Ile	Glu	Gly	Glu	Lys	Lys	Ala	Ala	Leu	Leu	Ala	Gln	Glu	Gly	
			2830					2835					2840		
acc	ccg	gcg	ctg	ggg	ctc	ccc	ggc	atc	agc	atg	gcc	cat	gat	gtt	62299
Thr	Pro	Ala	Leu	Gly	Leu	Pro	Gly	Ile	Ser	Met	Ala	His	Asp	Val	
			2845					2850					2855		
gag	gcc	cgg	cga	cga	agc	tgg	gat	gac	ctg	acg	gac	gag	tgg	gta	62344
Glu	Ala	Arg	Arg	Arg	Ser	Trp	Asp	Asp	Leu	Thr	Asp	Glu	Trp	Val	
			2860					2865					2870		
ccg	cac	ccc	gag	gcc	gca	tca	tac	ctt	gca	cag	ggc	cgc	acg	gtg	62389
Pro	His	Pro	Glu	Ala	Ala	Ser	Tyr	Leu	Ala	Gln	Gly	Arg	Thr	Val	
			2875					2880					2885		
ttc	gtc	tgc	ggc	gat	gcg	ccg	gac	atg	gac	ggc	acg	aac	atg	aac	62434
Phe	Val	Cys	Gly	Asp	Ala	Pro	Asp	Met	Asp	Gly	Thr	Asn	Met	Asn	
			2890					2895					2900		
gtc	gtt	cgt	gaa	gtt	gtg	cgg	atc	gcg	aag	ggc	ctt	tct	gcg	cgc	62479
Val	Val	Arg	Glu	Val	Val	Arg	Ile	Ala	Lys	Gly	Leu	Ser	Ala	Arg	
			2905					2910					2915		
gga	tgt	gac	cct	cag	att	ccg	tac	gtg	gac	cac	gac	cca	cag	cac	62524
Gly	Cys	Asp	Pro	Gln	Ile	Pro	Tyr	Val	Asp	His	Asp	Pro	Gln	His	
			2920					2925					2930		
gag	aaa	tcc	ggt	gtg	gat	gat	tgg	gtg	cat	cgc	gag	atg	aag	aag	62569
Glu	Lys	Ser	Gly	Val	Asp	Asp	Trp	Val	His	Arg	Glu	Met	Lys	Lys	
			2935					2940					2945		
ttg	ctt	gcc	gct	cat	cac	cac	aag	acg	gaa	gca	tgg	gct	cgt	cag	62614
Leu	Leu	Ala	Ala	His	His	His	Lys	Thr	Glu	Ala	Trp	Ala	Arg	Gln	
			2950					2955					2960		
ttc	ctg	atc	gtg	cag	tac	ttg	gcc	tct	gcg	cgc	ccg	gcg	ctt	cct	62659
Phe	Leu	Ile	Val	Gln	Tyr	Leu	Ala	Ser	Ala	Arg	Pro	Ala	Leu	Pro	
			2965					2970					2975		
tcg	gca	ctc	tgt	gac	tgg	ctc	gtg	gaa	cag	cat	gag	gcg	cac	ggt	62704
Ser	Ala	Leu	Cys	Asp	Trp	Leu	Val	Glu	Gln	His	Glu	Ala	His	Gly	
			2980					2985					2990		

179

ggg gtc aag gag atc gag atc aag cgc gct ctc gtc tgg tgc tcg	62749
Gly Val Lys Glu Ile Glu Ile Lys Arg Ala Leu Val Trp Cys Ser	
2995 3000 3005	
gtt tgg ttc tcc acc gag cat ggc atc ttc aag aag tgg gtc ggg	62794
Val Trp Phe Ser Thr Glu His Gly Ile Phe Lys Lys Trp Val Gly	
3010 3015 3020	
acc ttg aag agg aag aag gtc tgc aag ccc gag gtc atc gac cga	62839
Thr Leu Lys Arg Lys Lys Val Cys Lys Pro Glu Val Ile Asp Arg	
3025 3030 3035	
ttg gtc cag gcg ttg ccg agg gtc gag aac cgc ttc gag ccg cgc	62884
Leu Val Gln Ala Leu Pro Arg Val Glu Asn Arg Phe Glu Pro Arg	
3040 3045 3050	
aca tgg gtt caa cac tgg atg gag aag aac ggc gtc aca ttc gtg	62929
Thr Trp Val Gln His Trp Met Glu Lys Asn Gly Val Thr Phe Val	
3055 3060 3065	
tac gcg agc gat ggc ttc cag att ggg ggc cgg atg ttt cct gcg	62974
Tyr Ala Ser Asp Gly Phe Gln Ile Gly Gly Arg Met Phe Pro Ala	
3070 3075 3080	
tcc gag ttc ttg tcc cag atc gag atc gag tcc acg aac gag cgc	63019
Ser Glu Phe Leu Ser Gln Ile Glu Ile Glu Ser Thr Asn Glu Arg	
3085 3090 3095	
aca tgc atc agc gtg cga aat ctc ggt aac gcc gtg cgc ctg tgg	63064
Thr Cys Ile Ser Val Arg Asn Leu Gly Asn Ala Val Arg Leu Trp	
3100 3105 3110	
gag gcg gac caa cgc cgg ggg atc gtc gag cag tat cgt gag aag	63109
Glu Ala Asp Gln Arg Arg Gly Ile Val Glu Gln Tyr Arg Glu Lys	
3115 3120 3125	
ttg aag ttc cgc gag gac gga cgc cga ggt gac gcc gag gag ttc	63154
Leu Lys Phe Arg Glu Asp Gly Arg Arg Gly Asp Ala Glu Glu Phe	
3130 3135 3140	
att cga gca acc acg gga cgc ctt gac ccg gta gat ctc gtg gtc	63199
Ile Arg Ala Thr Thr Gly Arg Leu Asp Pro Val Asp Leu Val Val	
3145 3150 3155	
gtt ctt cac ttc att tgg cag gtg aag cgg aag ttg ttc ggg aag	63244
Val Leu His Phe Ile Trp Gln Val Lys Arg Lys Leu Phe Gly Lys	
3160 3165 3170	
ccg gtc aag gac cac ctg atg ccg atc ttc gtc ggc aaa cag cgg	63289
Pro Val Lys Asp His Leu Met Pro Ile Phe Val Gly Lys Gln Arg	
3175 3180 3185	
ggc ggg aag act acg gcc atc aag agc ctg cta cag cct ctt ggc	63334
Gly Gly Lys Thr Thr Ala Ile Lys Ser Leu Leu Gln Pro Leu Gly	
3190 3195 3200	
gat ctc gtt ggg ttc ccc gcc gat ctc cag ttc ctg acc gat gag	63379
Asp Leu Val Gly Phe Pro Ala Asp Leu Gln Phe Leu Thr Asp Glu	
3205 3210 3215	
cgc cag cag ttc cgg ttg act cgc tgc tac ggc atg ttc ttc gac	63424
Arg Gln Gln Phe Arg Leu Thr Arg Cys Tyr Gly Met Phe Phe Asp	
3220 3225 3230	

gaa atg gcg aag	gca gag cgg gta gcg	tct gac gtg ctc aag aac	63469
Glu Met Ala Lys	Ala Glu Arg Val Ala	Ser Asp Val Leu Lys Asn	
3235	3240	3245	
cgc atc acc gcc	gag cgg atc gag tac	cgg gtg ctc ggt gtc cag	63514
Arg Ile Thr Ala	Glu Arg Ile Glu Tyr	Arg Val Leu Gly Val Gln	
3250	3255	3260	
aag ctc cag acc	ggc acc aac aac gcc	acc ttc atc ggt gcg tcc	63559
Lys Leu Gln Thr	Gly Thr Asn Asn Ala	Thr Phe Ile Gly Ala Ser	
3265	3270	3275	
aac aag cag gtg	aag gag atc gtc gtc	gat ccc acg gga atg cga	63604
Asn Lys Gln Val	Lys Glu Ile Val Val	Asp Pro Thr Gly Met Arg	
3280	3285	3290	
cgc ttc tat caa	ctg gac tgt ctg ccc	cag atg aac tgg acg gca	63649
Arg Phe Tyr Gln	Leu Asp Cys Leu Pro	Gln Met Asn Trp Thr Ala	
3295	3300	3305	
ctg gag cag atc	gac ttc ctg gcc atg	tgg acg agc gtg gac gag	63694
Leu Glu Gln Ile	Asp Phe Leu Ala Met	Trp Thr Ser Val Asp Glu	
3310	3315	3320	
aag gcg gag agc	ccg ctc ctc ggc caa	ttg gcg cag gtc cag gcc	63739
Lys Ala Glu Ser	Pro Leu Leu Gly Gln	Leu Ala Gln Val Gln Ala	
3325	3330	3335	
cgc cag gat gag	att cga gcc aag gac	tcg gtg gaa gat ttc ttc	63784
Arg Gln Asp Glu	Ile Arg Ala Lys Asp	Ser Val Glu Asp Phe Phe	
3340	3345	3350	
gag cgc cgg tgc	gcg cag cgg gac cag	tgg atc aag gcc acg gcg	63829
Glu Arg Arg Cys	Ala Gln Arg Asp Gln	Trp Ile Lys Ala Thr Ala	
3355	3360	3365	
ctg tac ggc cac	tac gtc gag gac ttg	aag tac cag ggc cgt gac	63874
Leu Tyr Gly His	Tyr Val Glu Asp Leu	Lys Tyr Gln Gly Arg Asp	
3370	3375	3380	
cgg atg gcg ttc	tcg ctc acg aag ttc	ggt gaa cgc atg aag gag	63919
Arg Met Ala Phe	Ser Leu Thr Lys Phe	Gly Glu Arg Met Lys Glu	
3385	3390	3395	
ttg gcc ggt gaa	gag aac acg ggc tgg	aag cgg agc gat gga atc	63964
Leu Ala Gly Glu	Glu Asn Thr Gly Trp	Lys Arg Ser Asp Gly Ile	
3400	3405	3410	
aag tac cgg gtc	acg gtg aaa cgc tgg	gag acc gag atc gac ggg	64009
Lys Tyr Arg Val	Thr Val Lys Arg Trp	Glu Thr Glu Ile Asp Gly	
3415	3420	3425	
gtg atg gag agc	atc gcc gct ggc gct	cgg atg gac ctt gcg cac	64054
Val Met Glu Ser	Ile Ala Ala Gly Ala	Arg Met Asp Leu Ala His	
3430	3435	3440	
ccg tga ttggaacgga	tgcggtccga	ctgcggcgat gccttccatg	64110
Pro		ctgtctgctt	
cggtctcccc	ctggttcag	agagaggtgg	64170
		aaggatcgga agggatggaa	
		ggctatgggg	

aagtcaggac gatgacgccg cagcacgagt aaagggctga agtagccggt gctgtgaccc 64230
ttccggggcca tccatggcgt ccaacggggc tcaggtgccg ggccaagcgc ctcccgttcg 64290
ccttcaccgg cacggaggcg ggcagggcga gcgcccggaa ggacctcgga gagattcgat 64350
acattcagaa tagggacgga ctggtggcca ccggctcgat ccacaccttc gacaccatcc 64410
acggcctccg acggagaccg gttgagtaca agtccatgat cgagaacttc agctatatcc 64470
gtacaggctc cttgcttgct cggccggagc caccgttgga cctcattcct attcccattc 64530
cagtctttcg acctgccgtg ccattgagg accacgagcc cgcacccgtg gaatgcgcca 64590
tcctccagga actggccgag atcaagcgcg aggttcgtgc gcttcgcgct caatcgggtca 64650
cacagaacgc ctcggtgtg atgctggagc gcgcgtcggc aatgctcggc tgcaagcgggt 64710
cacaggctct caagctcctg gagcaaggcc ggttgaggcg ggcaccgaag atcggacgggt 64770
ccgtgatgat cagggtggcg agcatcgagg cactgctcgc ggagggcgcc cctcggaagg 64830
gaggcaatcg gcctcaaaa accaagcagc ctgaacgagt ccgcaccaag gagcggcggtg 64890
ggaagcagga gggaaaccgca gccggagcgt atccgaaccc aggtgttgcc atccgcaagc 64950
tcacgatttg agcctcggtg aaagtcgccg taacgaagcg gccgaagctc acgagcgagg 65010
gttttggtgc ccctcgacat ggcccaacgg gatagaggag gcagaacttc ttgtcctggc 65070
atccgcagcc ggatcgacac gctctaacc cccgacatcg tgacaagaat gatccggcag 65130
tgatcgaggc gtagttcgcc gcctcctact cctcgatcaa ggaagagact ttgctggccg 65190
cgcccagagg gccgatggtc tacttggcgc gcagggtcag cgtgtctcct gaccactcca 65250
ggcgcggtcg gtactgcgc tccaactccg tggtgagcac cttcgtcaac tcttccaagt 65310
gctccctatc gcagggtccg ttcgggtaca ggtccgccgc aacgatgttg atggcgctga 65370
tcggaatgga cgcgggaata gtcccgacgg ggtgagtggg cgtggaaggg agcagcgagc 65430
gcagcacgcc tccgaactcc gcaaccgtga tgggggctct ggtcggcttg gctccggcaa 65490
ttgcggcgcg agtctccggg aaggccgcct cgacgaacgc gatcgcgacc ctccgtcgcg 65550
tttctagatc ccacacccca aagttgaagg gcgcctgcat gagggcgcg ccatggagat 65610
cggttacgac tcctgcactg ccgttgagca tcgaggacac gcgcaggtgg aggtaccgga 65670
caacggtgcc cgccacgtcg gtgatggtgc acttcccccc gaaactagga ggttcggtta 65730
cggccaggat ctgcatcgaa tggtccttt gtgggaggcg tgcggtaatc acctggagcg 65790
ttcggcgcaa taagaacgtt gaaggctccg gcctggggcc tgacgcatca acgaacctcc 65850
tcgtactgat ctccgaactc tctggttcat gagcctcacc ccgcaggtgc ccgagctatg 65910
cgcaactccg atggatcatt cgaatgctcc aggcggatgg ggacgacgta catcggcggc 65970
actgtggcgc tgacgtagtg gagggccgtg gtccgcgtac tgctgtggtt gagcgccctga 66030
gcgacggctt ccagggacag gccaccgccc ttcgggttgt agatgctgcc ctgctcgacc 66090

agccacgttg cgaagctgtg ccggatcgcg ccgaagtggga cgaactccag gctcgtgcgc 66150
gcgcacacgc gctccacggc ctgcgcatg gtctgccgat ccggcgctga accccgggct 66210
tgaaggcgct gggcgggcgc gagcccctgc acatcgagcg acaagggtgaa cggcttcccg 66270
gtcttgtgct tgaacgtgac ggtggcggcg atctcgccct ggtccttcag caccgcgatc 66330
ttgccgtcac cgctcgcgag ccggtccacc tcggtgccgt gcattccggg cttggcctga 66390
aggcacagca catcgcgcat ggcctgcgag ccgagggccg cgtagagctg ctggacgtgc 66450
tcgatgctgt agcccttgag ttcgtgggtc cggcgggcag gggggacctt caggaagcgt 66510
ccgggggttca ccatcgggtc cagttcgccc gagtcctgaa ggaacgaggt gaaggacttg 66570
aagacgatga tctccactt cttgccggtg tcccagggtg ccaggatctt cttcaacgcc 66630
ggtccggtca ccgcgcgcag gtccttgtcc gcgagcgccg ttgcccacgc ggccaggtag 66690
ctgcgcgtgc tcttcaggta ctgcgcggac ctgccgctgg acttcagatg ctggaggaa 66750
cgccgggcac gatcctcgtc caggaacacc gcctgcgcgg cgtcctcctg ggcctgttgg 66810
gcctcgtctt ggcgctcctg gctcgcggtc cggtagcgcg ccgggttgcg gcggaaggcc 66870
gccagctcag ccagggcctc ggctcgcgc cgcacgtcca gcgtgacgtt gtacgcgacc 66930
ccggcgacgg tcttgcgcat gatccagcgc gtgcttccgt cgggtggcgt gtacgtgcgg 66990
cctccgatcc acttcccgt cacttcgtc atgcgcggct cccgtacccc cctcacttgg 67050
gtaaaacttg ggtacgagag gtacaggcgg gtggggcgtc ctgcaaccgc atcagggccg 67110
cgtcagaaaa cccaaggaa aacatggcgg tagagcatga gaatcgaact caccagggac 67170
gcctctcagc gcccctcacc ggttttgaag accggggcgg ccaccaggtc cggaggccct 67230
accaccgtcg attactgccg caccacagg gcccgctcaa ccaacgccc gctcacctg 67290
tcttcgcgc cgccttcacg ggcttcaacc ccggcaatcc gttcaccatc gtggacacca 67350
acagccgcgc cagctcctcc ggcggcgagc ccttgaagcc cgtacacgtc agccgcagcg 67410
ccacgatgcc gtgcacccc gcccacagca cctcgccag cttcgagggc tcggcgctct 67470
cggcaatcct ccccgccgc ttcaggctct cgaacacctt caccagcacc ccgaaggacc 67530
gcggccccgc gccgtcctcc ggcgcgctcc tgaacagcgc cgtcgacagc ttcgggtcct 67590
ccatgaagat gagccgttac gtctccggct gctccagccc gaacttcacg taggcctcgc 67650
ccagccggta cagccgctcc accggctcct ccacgggatg cgcgggctcg aacgcctcca 67710
gcaggctcct gaacccccgg atgcacagct cccgggcaat cgcgtcccgg ttctcgaagt 67770
ccaggtagag cgtcgccggc gcgtactcca ccgcgtccgc cagcttgcgc atcgagaggg 67830
caccaaaagc ctctcgctc accatgtcct tcgccacctg gaggatctgc tcgcgcagct 67890
ccgctcgcgt gcgctcctc cgtccgaaa tccccatct cccactgtag ggcttgacgt 67950

cctgaacggc cagcagtaga tgaacaccgt tcacaatgaa cgacgttcat aaaacgctcg 68010

ccgttcggaa cccagggagc ggacc atg gat acg gac aag gtg gcg ttg ttt 68062
Met Asp Thr Asp Lys Val Ala Leu Phe
3445 3450

gga gcg tcg ggc ccc atc gga cac agc atc gtg aag gct ctt cgc 68107
Gly Ala Ser Gly Pro Ile Gly His Ser Ile Val Lys Ala Leu Arg
3455 3460 3465

gag cag ggg agg gcg tac cgg gtg gtt ggc cgc tcg cgc gcc tcg 68152
Glu Gln Gly Arg Ala Tyr Arg Val Val Gly Arg Ser Arg Ala Ser
3470 3475 3480

ctt cag tcg cag ttc ggc gcg gac ccg aag gcg gag atc gcc acc 68197
Leu Gln Ser Gln Phe Gly Ala Asp Pro Lys Ala Glu Ile Ala Thr
3485 3490 3495

tgg aac ccc gac gcg ccg gac acc atc cgc gcc gcc gcc cgg ggc 68242
Trp Asn Pro Asp Ala Pro Asp Thr Ile Arg Ala Ala Ala Arg Gly
3500 3505 3510

atc cgc acg ctc atc tac atg gtc ggc gtg gac tac tgg cgc ttc 68287
Ile Arg Thr Leu Ile Tyr Met Val Gly Val Asp Tyr Trp Arg Phe
3515 3520 3525

cag ctc cac ccg gag ctc atg cgc cgc acc gtg gac gcc gcc atc 68332
Gln Leu His Pro Glu Leu Met Arg Arg Thr Val Asp Ala Ala Ile
3530 3535 3540

gcc gag ggc gtc gag cgc gtg gtg ctc atc ggc acc gtg tac ccg 68377
Ala Glu Gly Val Glu Arg Val Val Leu Ile Gly Thr Val Tyr Pro
3545 3550 3555

tac ggc cgc ccc cgc acc acg ccc gtc acc gag gac cac ccc cgc 68422
Tyr Gly Arg Pro Arg Thr Thr Pro Val Thr Glu Asp His Pro Arg
3560 3565 3570

gag ccg cac acc ttc aag ggc cgc atg cgc aag gag cag gag gac 68467
Glu Pro His Thr Phe Lys Gly Arg Met Arg Lys Glu Gln Glu Asp
3575 3580 3585

atc ctc ctc gcc gcg gac gcc gcc gga aag ctc cgc gcc acc atc 68512
Ile Leu Leu Ala Ala Asp Ala Ala Gly Lys Leu Arg Ala Thr Ile
3590 3595 3600

ctc cgc ctc ccg gac ttc tac ggc ccc ggc atc gac aag agc ttc 68557
Leu Arg Leu Pro Asp Phe Tyr Gly Pro Gly Ile Asp Lys Ser Phe
3605 3610 3615

ctc cac tcc gcc ttc gtc gcc gcc gcc cag ggc aag cgc gcc cag 68602
Leu His Ser Ala Phe Val Ala Ala Ala Gln Gly Lys Arg Ala Gln
3620 3625 3630

ctc atc ggc ccc atc gac ccg ccg cac gag tac gtc tac gtg ccc 68647
Leu Ile Gly Pro Ile Asp Pro Pro His Glu Tyr Val Tyr Val Pro
3635 3640 3645

gac gtg ggt ccc acc gtg acg gcc ctc atg gac gcc ccc ggt gcc 68692
Asp Val Gly Pro Thr Val Thr Ala Leu Met Asp Ala Pro Gly Ala
3650 3655 3660

tac ggc cgc tgg tgg aac ctc gcc ggc gcc ggc gtc acc tca cag 68737

[illegible]

tccacgccac ccaccccctc tcgcttcgcg gtcgcttcc ccatgggccg gcacatatga	70074
aggcacgctg tgcgttgta acgacgggca cggctgttag gttccggcgc ccgtccatgt	70134
gtgacggccc gccccgtgaa catccttgtc gtcgatgacg atctcgagct gtgcaccatg	70194
ctctctcgct tcctggagat gcatggctac acggtctact cggcggcgga tgcctgcag	70254
gcgctcgaca tcctggagcg caaccagggtg ggcatgggtca tcaccgatta cctcatgccc	70314
cacctggacg gcatccactt cacggagatg ctcaaggccg acccccgctt ccaggccatt	70374
cccgtgctgc tgatgacggg cgcccaggac gagggcggtca ccgaccgggg cctgcgcaag	70434
ggcgtggcgc tgacgctgcg caaacccctg gacatgggtc agctgctcaa cctcgtgcgc	70494
ttcgccgagt agccccccgg ccagacggcc ccgagggggc ctggccccgt ccgctcgccc	70554
acaggcacgc ccttcctggt tgacatcatt tcgctggccc ttatgttggc gctcgccatg	70614
gccgagtgtt aacggccagc gtcgtcacag ttaaattccc cagtaatttc aggaggttgc	70674
g atg gca gcg aag gag att ttc ttc cac cag tcc gcg cgt gag gcc	70720
Met Ala Ala Lys Glu Ile Phe Phe His Gln Ser Ala Arg Glu Ala	
3775 3780 3785	
atc ctt cgc ggt gtc cgg acc ctc gcg gat gcc gtg gcg gtg acc	70765
Ile Leu Arg Gly Val Arg Thr Leu Ala Asp Ala Val Ala Val Thr	
3790 3795 3800	
ctg ggc ccc aag ggc cgc aac gtg gtc atc gag aag agc ttt ggt	70810
Leu Gly Pro Lys Gly Arg Asn Val Val Ile Glu Lys Ser Phe Gly	
3805 3810 3815	
tcg ccc acg atc acc aag gac ggc gtc acc gtc gcc aag gag atc	70855
Ser Pro Thr Ile Thr Lys Asp Gly Val Thr Val Ala Lys Glu Ile	
3820 3825 3830	
gac ctc gac aac aag ttc gag aac atg ggc gcg cag atg gtg aag	70900
Asp Leu Asp Asn Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys	
3835 3840 3845	
gag gtc gcg tcg aag acc tcc gac aag gcc ggc gac ggc acc acg	70945
Glu Val Ala Ser Lys Thr Ser Asp Lys Ala Gly Asp Gly Thr Thr	
3850 3855 3860	
acc gcc acg gtg ctg gcg cgt gcc atc tac gag gag ggc ctc aag	70990
Thr Ala Thr Val Leu Ala Arg Ala Ile Tyr Glu Glu Gly Leu Lys	
3865 3870 3875	
ctg gtg gcc gcc ggc cac agc ccc atg gac ctc aag cgc ggc atc	71035
Leu Val Ala Ala Gly His Ser Pro Met Asp Leu Lys Arg Gly Ile	
3880 3885 3890	
gac aag gcc gtc gag gtc gtc gtg gag gag ctg aag aag ctg tcc	71080
Asp Lys Ala Val Glu Val Val Val Glu Glu Leu Lys Lys Leu Ser	
3895 3900 3905	
aag ccc acg gcg gac aag aag gcc atc acc cag gtg ggc acc atc	71125
Lys Pro Thr Ala Asp Lys Lys Ala Ile Thr Gln Val Gly Thr Ile	
3910 3915 3920	
tcc gcc aac ggt gat gag acc atc ggc tcc atc atc gcg gac gcg	71170

186

Ser 3925	Ala	Asn	Gly	Asp	Glu 3930	Thr	Ile	Gly	Ser	Ile 3935	Ile	Ala	Asp	Ala	
atg	gag	aag	gtg	ggc	aag	gag	ggc	gtc	atc	acc	gtc	gag	gag	gcc	71215
Met 3940	Glu	Lys	Val	Gly	Lys 3945	Glu	Gly	Val	Ile	Thr 3950	Val	Glu	Glu	Ala	
aag	ggc	ctg	gag	acc	aac	ctc	gac	gtg	gtg	gag	ggc	atg	cag	ttc	71260
Lys 3955	Gly	Leu	Glu	Thr	Asn 3960	Leu	Asp	Val	Val	Glu 3965	Gly	Met	Gln	Phe	
gac	cgc	ggc	tac	gtc	tct	ccg	tac	ttc	gtg	acg	aac	cgc	gag	cgc	71305
Asp 3970	Arg	Gly	Tyr	Val	Ser 3975	Pro	Tyr	Phe	Val	Thr 3980	Asn	Arg	Glu	Arg	
atg	gag	gtc	gtc	atg	gac	gac	ccc	ttc	atc	ctc	atc	agc	gag	aag	71350
Met 3985	Glu	Val	Val	Met	Asp 3990	Asp	Pro	Phe	Ile	Leu 3995	Ile	Ser	Glu	Lys	
aag	gtc	tcg	tcg	atg	cag	gac	atg	att	ccc	atc	ctg	gag	cag	gtg	71395
Lys 4000	Val	Ser	Ser	Met	Gln 4005	Asp	Met	Ile	Pro	Ile 4010	Leu	Glu	Gln	Val	
gcg	cgc	tcc	ggg	aag	ccg	ctg	ctc	atc	atc	gcc	gac	gac	atc	gag	71440
Ala 4015	Arg	Ser	Gly	Lys	Pro 4020	Leu	Leu	Ile	Ile	Ala 4025	Asp	Asp	Ile	Glu	
ggc	gag	gcg	ctg	gcc	acc	ctg	gtg	gtc	aac	aag	atc	cgc	ggc	gtg	71485
Gly 4030	Glu	Ala	Leu	Ala	Thr 4035	Leu	Val	Val	Asn	Lys 4040	Ile	Arg	Gly	Val	
ctg	aac	gtg	gcc	gcg	gtg	aag	gcg	ccg	ggc	ttc	ggt	gac	cgc	cgc	71530
Leu 4045	Asn	Val	Ala	Ala	Val 4050	Lys	Ala	Pro	Gly	Phe 4055	Gly	Asp	Arg	Arg	
aag	gag	atg	ctg	aag	gac	atc	gcc	acc	ctg	acg	ggc	ggc	atg	gtc	71575
Lys 4060	Glu	Met	Leu	Lys	Asp 4065	Ile	Ala	Thr	Leu	Thr 4070	Gly	Gly	Met	Val	
gtc	agc	gag	gag	ctg	ggc	cac	aag	tac	gag	aac	ctg	acg	ctc	aac	71620
Val 4075	Ser	Glu	Glu	Leu	Gly 4080	His	Lys	Tyr	Glu	Asn 4085	Leu	Thr	Leu	Asn	
gac	ctg	ggc	cgc	tcc	aag	cgc	atc	acg	gtg	gac	aag	gac	aac	acc	71665
Asp 4090	Leu	Gly	Arg	Ser	Lys 4095	Arg	Ile	Thr	Val	Asp 4100	Lys	Asp	Asn	Thr	
acc	atc	gtc	gac	ggc	gcc	ggc	acc	aag	agc	gag	atc	gag	ggc	cgc	71710
Thr 4105	Ile	Val	Asp	Gly	Ala 4110	Gly	Thr	Lys	Ser	Glu 4115	Ile	Glu	Gly	Arg	
atc	aag	ctc	atc	cgc	tcg	cag	atc	gac	acc	gtc	acc	agc	gac	tac	71755
Ile 4120	Lys	Leu	Ile	Arg	Ser 4125	Gln	Ile	Asp	Thr	Val 4130	Thr	Ser	Asp	Tyr	
gac	cgc	gag	aag	ctc	cag	gag	cgg	ctg	gcc	aag	ctg	gtg	ggc	ggc	71800
Asp 4135	Arg	Glu	Lys	Leu	Gln 4140	Glu	Arg	Leu	Ala	Lys 4145	Leu	Val	Gly	Gly	
gtg	gcc	gtc	atc	aac	gtg	ggc	gcc	gcg	acc	gag	acc	gag	atg	aag	71845
Val 4150	Ala	Val	Ile	Asn	Val 4155	Gly	Ala	Ala	Thr	Glu 4160	Thr	Glu	Met	Lys	

gcaccaggga gccgccggga ccggtgtaga gaacctggga gttgccgcca tcggtgtcca 72999
cccggaccag cgactgccgc cccgtagcga ccaccagcac gtcgcgtccg gcgagcgtgg 73059
cgcgagcac ctcggtacca ccgacggtga cgttgcgatc cgtcgcgtcg atgggcaggt 73119
cgccccgctt cccggttgcc ttgtccaccc gccagagccc gtccaggtcg agcacgtaga 73179
cgctcgtccc atccacggcg atggcgtctg gtgcgcggaa gcccgtaggc agaaccgtgg 73239
gtggacctcc cgcgcggggc agcctcagaa cctgtcccgg tcccggactc ggctgctggg 73299
ggttcaggga gt gtg act cgg aga tgt aga tgt ccg tcg tgt cca ccg 73347
Val Thr Arg Arg Cys Arg Cys Pro Ser Cys Pro Pro
4325 4330

cga gcc ggc ggg gcg tgt tca gtc gcg tgg cca gct ccg tat cgg 73392
Arg Ala Gly Gly Ala Cys Ser Val Ala Trp Pro Ala Pro Tyr Arg
4335 4340 4345

gct cgg ccg cgg gcg cct cac agc ttc cgc cat cca ccg tgc ccg 73437
Ala Arg Pro Arg Ala Pro His Ser Phe Arg His Pro Pro Cys Pro
4350 4355 4360

cgt cgc ccg ctc ccg cat ccg tgc ccg cat cgg agc cgg agc cac 73482
Arg Arg Pro Leu Pro His Pro Cys Pro His Arg Ser Arg Ser His
4365 4370 4375

cat ccg gct ccg tgc ctc cgt cct ccg tcg ggc ccg cgt cct ctc 73527
His Pro Ala Pro Cys Leu Arg Pro Pro Ser Gly Pro Arg Pro Leu
4380 4385 4390

cgc ccg tcc cgg agt cct ccg tcg cgc ccc cgt ccg gcc gct ccg 73572
Arg Pro Ser Arg Ser Pro Pro Ser Arg Pro Arg Pro Ala Ala Pro
4395 4400 4405

tcg gcg ggg gct tgt ccg agg aac agg cca ccg cga aga gca gga 73617
Ser Ala Gly Ala Cys Pro Arg Asn Arg Pro Pro Arg Arg Ala Gly
4410 4415 4420

cgg aca gga gcg cga gtc tca tgg gac ctc tct ctt gaa ggc gcg 73662
Arg Thr Gly Ala Arg Val Ser Trp Asp Leu Ser Leu Glu Gly Ala
4425 4430 4435

gct cag ctc acg gct gag cgc cgt ggt atc gga cag cag ctt ccg 73707
Ala Gln Leu Thr Ala Glu Arg Arg Gly Ile Gly Gln Gln Leu Arg
4440 4445 4450

aag gca cgc ggc ggc acc ggc gcg cag cga ctc cag cgc cgt ctc 73752
Lys Ala Arg Gly Gly Thr Gly Ala Gln Arg Leu Gln Arg Arg Leu
4455 4460 4465

cat ggt gag gtg ctc ggc cag cac gac gaa ggg ggc gcc ctg ggc 73797
His Gly Glu Val Leu Gly Gln His Asp Glu Gly Gly Ala Leu Gly
4470 4475 4480

cag cgc ctt ggc cag ctc cag cgc ctt gcg ccc ata ggc cgg cgc 73842
Arg Arg Leu Gly Gln Leu Gln Arg Leu Ala Pro Ile Gly Arg Arg
4485 4490 4495

gaa gca cca gct cac cac cac gcc cac cgg cgt cgc cat cgc cat 73887
Glu Val Pro Ala His His His Ala His Arg Arg Arg His Arg His

4500	4505	4510	
cag ctc cgt cga ggg cag cac ccg cgc ctc cag ccc cag cat gtc Gln Leu Arg Arg Gly Gln His Pro Arg Leu Gln Pro Gln His Val 4515 4520 4525			73932
cag ggc ctc gga gat gag ccg cgc cgt cgt cgg gtt gtc ctc cag Gln Gly Leu Gly Asp Glu Pro Arg Arg Arg Arg Val Val Leu Gln 4530 4535 4540			73977
cac gtc cac ccg ccg cac ttc ggc ccc ggg ggc cac cac cgg agg His Val His Pro Pro His Phe Gly Pro Gly Gly His His Arg Arg 4545 4550 4555			74022
cac cgg cgg ctc cga gga cag cgt cgt gcg caa cag ccc gcg cac His Arg Arg Leu Arg Gly Gln Arg Arg Ala Gln Gln Pro Ala His 4560 4565 4570			74067
ctg ccg gat gtc gtc gaa ggg ctt gag cag gta gtc cac cac gcc Leu Pro Asp Val Val Glu Gly Leu Glu Gln Val Val His His Ala 4575 4580 4585			74112
cag ctc cag cgc ctg ctg cgt cgt cac cag cga cgg gta gcc cgt Gln Leu Gln Arg Leu Leu Arg Arg His Gln Arg Arg Val Ala Arg 4590 4595 4600			74157
cat cag gat gac gcg cga gtt gga gta gag ccg ccg cgc ctg ctg His Gln Asp Asp Ala Arg Val Gly Val Glu Pro Pro Arg Leu Leu 4605 4610 4615			74202
cgc cag ctc cag ccc gga cag gcc cgg cag gtt ctt gtc cgt gac Arg Gln Leu Gln Pro Gly Gln Ala Arg Gln Val Leu Val Arg Asp 4620 4625 4630			74247
gat gag gtc cac tgg cgc ctg gcg cag caa atc cag cgc ttc ctc Asp Glu Val His Trp Arg Leu Ala Gln Gln Ile Gln Arg Phe Leu 4635 4640 4645			74292
gcc gct ggc cgc ctc gat gac ctc gca ctc ctt gcc cat cag gtc Ala Ala Gly Arg Leu Asp Asp Leu Ala Leu Leu Ala His Gln Val 4650 4655 4660			74337
gcg gaa gac cat gcg gat gat ggt ctc gtc gtc cac cac cag cag Ala Glu Asp His Ala Asp Asp Gly Leu Val Val His His Gln Gln 4665 4670 4675			74382
ccg ctt gcg cag ctg cgg cgg cgc gtc cgt ggc ggg gaa gag tac Pro Leu Ala Gln Leu Arg Arg Arg Val Arg Gly Gly Glu Glu Tyr 4680 4685 4690			74427
ccg gaa cac ggt ggc cgg cgg ggg cac gtc gcg aat ggc gcc cgg Pro Glu His Gly Gly Arg Arg Gly His Val Ala Asn Gly Ala Arg 4695 4700 4705			74472
cgc cgc gag gcc aat cgt cgc gtt gtg ctc ctg ggc gat gcg ctt Arg Arg Glu Ala Asn Arg Arg Val Val Leu Leu Gly Asp Ala Leu 4710 4715 4720			74517
gca cac cgc cag ccc cag tcc cgt gcc gcg ctt gtt ggc ggt gac Ala His Arg Gln Pro Gln Ser Arg Ala Ala Leu Val Gly Gly Asp 4725 4730 4735			74562
gta ggg ctc gaa gat gcg ctc gcg cag ctc cac ggg gat gcc cgg			74607

Val Gly 4740	Leu Glu Asp 4745	Ala Leu 4745	Ala Gln Leu His 4750	Gly Asp Ala Arg 4750	
gcc cca Ala Pro 4755	gtc cgc cac gta Val Arg His Val 4760	cag Gln 4760	cac cgg aga ggc His Arg Arg Gly 4765	gcc ctc gcg cga Ala Leu Ala Arg 4765	74652
gag cac Glu His 4770	cac ctt cac gcg His Leu His Ala 4775	ccc Pro 4775	gcg gcc ggc cat Ala Ala Gly His 4780	ggc gtc gcg cgc Gly Val Ala Arg 4780	74697
gtt gtt Val Val 4785	cag cag gtt cag Gln Gln Val Gln 4790	cgt Arg 4790	gag ctg ctc gat Glu Leu Leu Asp 4795	gag ccg cgc gtt Glu Pro Arg Val 4795	74742
gcc ctg Ala Leu 4800	gat ggt gat gtc Asp Gly Asp Val 4805	ctc Leu 4805	ggg cgc ctc cac Gly Arg Leu His 4810	ctc cag aga cat Leu Gln Arg His 4810	74787
gcg cgc Ala Arg 4815	cga gtc cgg gtt Arg Val Arg Val 4820	gac Asp 4820	gct gaa cag ctt Ala Glu Gln Leu 4825	cgc cgc cgc cca Arg Arg Arg Pro 4825	74832
gat ggg Asp Gly 4830	cgc cgc cag cga Arg Arg Gln Arg 4835	cag Gln 4835	gcg ctg ctg cgg Ala Leu Leu Arg 4840	cgc gtg ccg ctc Arg Val Pro Leu 4840	74877
gct ggc Ala Gly 4845	gag gcg gat gta Glu Ala Asp Val 4850	gtc Val 4850	gga gac gat ctg Gly Asp Asp Leu 4855	ttc cat ccg ctc Phe His Pro Leu 4855	74922
cac ctg His Leu 4860	cgc cag cag cag Arg Gln Gln Gln 4865	ccg Pro 4865	cag cgg gcc ggt Gln Arg Ala Gly 4870	ggg gcc acc gtc Gly Ala Thr Val 4870	74967
ctc cgc Leu Arg 4875	gag cag ctg cgc Glu Gln Leu Arg 4880	gta Val 4880	ggc ctt cac gcc Gly Leu His Ala 4885	cag ggg ctg Gln Gln Gly Leu 4885	75012
ccg cag Pro Gln 4890	ctc gtg cag cac Leu Val Gln His 4895	ctc Leu 4895	cgc cgc cag ctg Arg Arg Gln Leu 4900	tgc cga gtg cgc Cys Arg Val Arg 4900	75057
gcc cgc Ala Arg 4905	acg ttg cag agc Thr Leu Gln Ser 4910	cgc Arg 4910	cgc cgc cgc gcg Arg Arg Arg Ala 4915	ggc cgg cag Gly Arg Gln 4915	75102
atc tcc Ile Ser 4920	cgc ctc cag ggc Arg Leu Gln Gly 4925	ctg Leu 4925	gag caa ctg gga Glu Gln Leu Gly 4930	gag cgg cgc ggg Glu Arg Arg Gly 4930	75147
tgt ttc Cys Phe 4935	cat ggc ccg agc His Gly Pro Ser 4940	atg Met 4940	gcg gag gca gcg Ala Glu Ala Ala 4945	gta tct acg caa Val Ser Thr Gln 4945	75192
cgc ggg Arg Gly 4950	cgg agg gat tcc Arg Arg Asp Ser 4955	gtt Val 4955	gac ccg ggc ccc Asp Pro Gly Pro 4960	tgg ccg cgc gga Trp Pro Arg Gly 4960	75237
tgc tcc Cys Ser 4965	tcg cgt ccc ctg Ser Arg Pro Leu 4970	gcg Ala 4970	ccc cgt cca gaa Pro Arg Pro Glu 4975	cgt gga cgc cac Arg Gly Arg His 4975	75282

cac cgg cct cgg gag ggc ctt cat gcc gca gac caa ccc gtt cca	75327
His Arg Pro Arg Glu Gly Leu His Ala Ala Asp Gln Pro Val Pro	
4980 4985 4990	
ctc gct ggt gcc ccg gaa gat gac cga ctc cga gct ggc ccg ctc	75372
Leu Ala Gly Ala Pro Glu Asp Asp Arg Leu Arg Ala Gly Pro Leu	
4995 5000 5005	
cat ccg gct gaa cat cga ggc gga gct gga cgc cat caa cct cta	75417
His Pro Ala Glu His Arg Gly Gly Ala Gly Arg His Gln Pro Leu	
5010 5015 5020	
cgc cgc cca cat cga tgc cac cga caa cga gga ggc caa ggc cat	75462
Arg Arg Pro His Arg Cys His Arg Gln Arg Gly Gly Gln Gly His	
5025 5030 5035	
cct ccg tca cgt cat gga cga gga gcg cga gca cgc cgc cct ctt	75507
Pro Pro Ser Arg His Gly Arg Gly Ala Arg Ala Arg Arg Pro Leu	
5040 5045 5050	
ctg gca gct gat tgc ccg gtt gga ccc cga gca ggc cca gca tga	75552
Leu Ala Ala Asp Cys Pro Val Gly Pro Arg Ala Gly Pro Ala	
5055 5060 5065	
ccgcgaggcc tcggagaagt accggctcat caccaccggt gccccgcacg aggaggtgga	75612
ggcgggtgggc aaggagggcg ggggagcacg cgccgaggtg gagaccaccc cgaccaagcg	75672
cctcaccgtg ggcaacctgc gccgctgagc cgctacgtgc tggagccgcc ggggccgctg	75732
ccgccctggc gccgctccgc cggcagcggc tccagatcca gggccgccat gtagacgaca	75792
ttccgggagc cacgcttgcc ccggtacatg gcccggtccg ccaggtccag cagcgtggcc	75852
ttgtcccggg cgtgctcggg gaagctggcc acgccgatgc acgtgctgag cttcagcgac	75912
agccccctgc gcgcgaggaa caggtgcgtc tccatggtgc ggcggatgc ctccgccacc	75972
ttgagcgcg cgcgggagtc ggtgttgcg aggagcacca cgtactcgtc gccgccgtag	76032
cgggccacca cgtcgtggtc tctcacgcag cccttcacca cccgggccgt ctccaccaac	76092
agccgcgagc ccacgaggtg gccgtgcgtg tcgttgatgg acttgaagtg gtccaggtcc	76152
aggaacagca ggctgaagg ggcgtgcgtc tggacggagt cctggacctc gcggtccagc	76212
accaggtgca ggtagcgcgt gttgaagagg cgggtgaggt cgtcgacgta cgccaggtcc	76272
tccaccgccg cgaagcgccc gaggttgcg agcgccagtg ccagttgcy caccaggtag	76332
cccgccgtct cgcgggtcca ctccggctcg gcgcgctcga agaagagcac cgcgtggccc	76392
agcaccgtgg tgccctccac ggcggggatg gagaggacgc gcgggaaggc ggcgtccagc	76452
ccttccagct cccggggcg gcgcgcgttc gagaggcgct cgacgagggg gccaccagc	76512
tccggctcgt gccgcggagg cagtccgcgg acgccctgga gccgcaggcc cttgtcgccg	76572
tcgcgctcca gcagcaccac ggcgctggcg caggccatgg actccagcg gctcgtggtg	76632
gcggaggcca gcttctcgcg ctccagcgtg gtggcgatgc gctggccggc ctccagcatg	76692
gcgacgtgcc ggcgacgca ggcgttctcc tggaggaggt ctgcgctggt gagggcacgg	76752

ctg	ctg	gcg	ggg	acc	ttc	gag	gcc	ttc	gac	gag	aac	cgc	gcc	cgc	78363
Leu	Leu	Ala	Gly	Thr	Phe	Glu	Ala	Phe	Asp	Glu	Asn	Arg	Ala	Arg	
5120					5125					5130					
gcc	gcc	cgc	gcc	gat	gac	cac	atc	gcc	aag	ccc	ttc	gag	agc	cag	78408
Ala	Ala	Arg	Ala	Asp	Asp	His	Ile	Ala	Lys	Pro	Phe	Glu	Ser	Gln	
5135					5140					5145					
gtg	ctg	ctc	gac	aag	gtg	aag	gcg	ctg	gtg	ggg	cag	aag	tcc	aac	78453
Val	Leu	Leu	Asp	Lys	Val	Lys	Ala	Leu	Val	Gly	Gln	Lys	Ser	Asn	
5150					5155					5160					
acc	atg	ccg	gcc	tcc	gcg	gcc	acg	cag	gtg	cgc	ccg	ccc	acc	gcc	78498
Thr	Met	Pro	Ala	Ser	Ala	Ala	Thr	Gln	Val	Arg	Pro	Pro	Thr	Ala	
5165					5170					5175					
cag	ccg	gcc	act	ccc	gcg	ccg	gtg	cct	ccg	ggc	ccg	cag	ccg	ggc	78543
Gln	Pro	Ala	Thr	Pro	Ala	Pro	Val	Pro	Pro	Gly	Pro	Gln	Pro	Gly	
5180					5185					5190					
gcg	ccc	acg	gcc	gcc	cgt	ccg	ggc	att	ccg	cag	ggc	acc	cag	ccg	78588
Ala	Pro	Thr	Ala	Ala	Arg	Pro	Gly	Ile	Pro	Gln	Gly	Thr	Gln	Pro	
5195					5200					5205					
ggc	gtg	ccg	cgt	cct	ccg	ggc	ccg	ggc	atg	ccg	ccg	ggt	gct	ccg	78633
Gly	Val	Pro	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Gly	Ala	Pro	
5210					5215					5220					
ggc	gcg	ccg	cgt	cct	ccg	ggg	cct	gga	atg	ccg	ccg	ggg	ccg	ggc	78678
Gly	Ala	Pro	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Gly	Pro	Gly	
5225					5230					5235					
atg	gcc	cgt	cct	ccc	ggg	cct	gga	atg	ccg	cct	ccg	ggt	gct	ccg	78723
Met	Ala	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Pro	Gly	Ala	Pro	
5240					5245					5250					
ggc	gcg	ccg	cgt	cct	ccg	ggg	cct	gga	atg	ccg	ccg	gga	ccg	ggc	78768
Gly	Ala	Pro	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Gly	Pro	Gly	
5255					5260					5265					
atg	gcc	cgc	ccg	ccg	ggg	cct	gga	atg	ccg	ccg	ccg	ggt	gct	ccg	78813
Met	Ala	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Pro	Gly	Ala	Pro	
5270					5275					5280					
ggc	gcg	ccg	cgt	cct	ccg	ggc	ccg	ggc	atg	ccg	cct	ccg	ggc	gcc	78858
Gly	Ala	Pro	Arg	Pro	Pro	Gly	Pro	Gly	Met	Pro	Pro	Pro	Gly	Ala	
5285					5290					5295					
gcc	gcg	cgt	ccc	ggg	cgt	gcc	gcc	tcc	gcc	ggg	tgc	gtc	cgc	gcc	78903
Ala	Ala	Arg	Pro	Gly	Arg	Ala	Ala	Ser	Ala	Gly	Cys	Val	Arg	Ala	
5300					5305					5310					
ggg	cat	gcc	gcc	gcg	tcc	ggg	cgt	gcc	cgc	gcc	tcc	ggg	tgg	gcc	78948
Gly	His	Ala	Ala	Ala	Ser	Gly	Arg	Ala	Arg	Ala	Ser	Gly	Trp	Ala	
5315					5320					5325					
ggc	gcc	ggc	atg	gcc	cgt	ccg	ggg	att	ccc	cag	ggc	agc	cag	ccg	78993
Gly	Ala	Gly	Met	Ala	Arg	Pro	Gly	Ile	Pro	Gln	Gly	Ser	Gln	Pro	
5330					5335					5340					
ggc	ttc	ccg	cgt	cct	ccg	ggc	ggc	gcc	gcg	ccg	ctg	cct	ccc	gcg	79038
Gly	Phe	Pro	Arg	Pro	Pro	Gly	Gly	Ala	Ala	Pro	Leu	Pro	Pro	Ala	
5345					5350					5355					

ccg ggt ggc gcg acc cag	ccg gcg gcc cgc gcg	agg gac ccg ttc	79083
Pro Gly Gly Ala Thr Gln	Pro Ala Ala Arg Ala	Arg Asp Pro Phe	
5360	5365	5370	
ggc ctg ggc gct ccc acg	gcg gcg ccg cag gca	gcc cag cct tcc	79128
Gly Leu Gly Ala Pro Thr	Ala Ala Pro Gln Ala	Ala Gln Pro Ser	
5375	5380	5385	
atc agc atc gag gac tcg	ctg ccc gag ccc agc	ggt gcc gag gag	79173
Ile Ser Ile Glu Asp Ser	Leu Pro Glu Pro Ser	Gly Ala Glu Glu	
5390	5395	5400	
atc tcc ctg gac atg ggc	ggc ccc ccg acg gcc	gct ccg acc gcg	79218
Ile Ser Leu Asp Met Gly	Gly Pro Pro Thr Ala	Ala Pro Thr Ala	
5405	5410	5415	
gtg gcg ccg acc gcc gct	ccg gcc acg gcc cgt	ccg acc gcc gcg	79263
Val Ala Pro Thr Ala Ala	Pro Ala Thr Ala Arg	Pro Thr Ala Ala	
5420	5425	5430	
gat ggt ggc gag gcg ctg	ctg cgc gag gcg ctc	tcg aag gcc tcc	79308
Asp Gly Gly Glu Ala Leu	Leu Arg Glu Ala Leu	Ser Lys Ala Ser	
5435	5440	5445	
cgc gag gtc atc gag aag	att gcc tgg gag gtg	gta ccg cag ctg	79353
Arg Glu Val Ile Glu Lys	Ile Ala Trp Glu Val	Val Pro Gln Leu	
5450	5455	5460	
gcg gag acc atc atc cgt	gag gag ctc gag cgg	ctc atc aag gac	79398
Ala Glu Thr Ile Ile Arg	Glu Glu Leu Glu Arg	Leu Ile Lys Asp	
5465	5470	5475	
cgg gag acg cag cac tga	cgcggttcctc gccttaacct	ctctccaaga	79446
Arg Glu Thr Gln His			
5480			
ccttcccgcc ggccccgggtt	cccccggggc tggccgtccg	caatgaccga caccactgaa	79506
ctgtccaagg cctacgagcc	caccgagggtc gagggccgct	ggtacgcctg ctgg atg	79563
		Met	
		5485	
gag cgg aac tac ttc	cgc gcc gag gcg acc	tcc gac aag ccc gcc	79608
Glu Arg Asn Tyr Phe	Arg Ala Glu Ala Thr	Ser Asp Lys Pro Ala	
5490	5495	5500	
ttc tcc atc gtg ctg	ccg ccg ccc aac gtg	acg ggc agc ctc cac	79653
Phe Ser Ile Val Leu	Pro Pro Pro Asn Val	Thr Gly Ser Leu His	
5505	5510	5515	
atc ggc cac gcg ctc	acc gcc acc atc cag	gac atc ctc gcg cgc	79698
Ile Gly His Ala Leu	Thr Ala Thr Ile Gln	Asp Ile Leu Ala Arg	
5520	5525	5530	
tgg aag cgg atg agc	ggc ttc aac acc ctg	tgg ctc ccc ggc acg	79743
Trn Lys Arg Met Ser	Gly Phe Asn Thr Leu	Trp Leu Pro Gly Thr	
5535	5540	5545	
gac cac gcc ggc atc	gcc acg cag atg gtg	gtg gag aag gag ctg	79788
Asp His Ala Gly Ile	Ala Thr Gln Met Val	Val Glu Lys Glu Leu	
5550	5555	5560	

195

aag aag acc gag ggc	aag agc cgg cac gac	ctc ggc cgc gcc gcc	79833
Lys Lys Thr Glu Gly	Lys Ser Arg His Asp	Leu Gly Arg Ala Ala	
5565	5570	5575	
ttc ctg gag cgc gtc	tgg gag tgg aag ggc	aag tac ggc gcc cgc	79878
Phe Leu Glu Arg Val	Trp Glu Trp Lys Gly	Lys Tyr Gly Ala Arg	
5580	5585	5590	
atc ggc gag cag cac	cgc ttc ctg ggc gcc	tcg ctg gac tgg agc	79923
Ile Gly Glu Gln His	Arg Phe Leu Gly Ala	Ser Leu Asp Trp Ser	
5595	5600	5605	
cgc gag cgc ttc acc	atg gac gag cag tcc	tcc gcc gcc gtg cgc	79968
Arg Glu Arg Phe Thr	Met Asp Glu Gln Ser	Ser Ala Ala Val Arg	
5610	5615	5620	
gag gtc ttc gtg cgc	ctg tac gaa gag ggc	ctg atg tac cgg gcc	80013
Glu Val Phe Val Arg	Leu Tyr Glu Glu Gly	Leu Met Tyr Arg Ala	
5625	5630	5635	
cag aag ctc atc aac	tgg tgc ccc tcg tgc	cgc acc gcc ctc agc	80058
Gln Lys Leu Ile Asn	Trp Cys Pro Ser Cys	Arg Thr Ala Leu Ser	
5640	5645	5650	
gac ttg gag gtg gag	cac gag gag tcg aag	aag ggc tcc atc tgg	80103
Asp Leu Glu Val Glu	His Glu Glu Ser Lys	Lys Gly Ser Ile Trp	
5655	5660	5665	
cac atc cgc tac ccc	gtc aag gac agc gac	cgc acg ctc acc gtg	80148
His Ile Arg Tyr Pro	Val Lys Asp Ser Asp	Arg Thr Leu Thr Val	
5670	5675	5680	
gcc acc aca cgc ccg	gag acg atg ctg ggc	gac acc gcg gtg gcc	80193
Ala Thr Thr Arg Pro	Glu Thr Met Leu Gly	Asp Thr Ala Val Ala	
5685	5690	5695	
atc cac ccg gag gac	gag cgc tac ctg ggg	ctg gcc ggc aag agc	80238
Ile His Pro Glu Asp	Glu Arg Tyr Leu Gly	Leu Ala Gly Lys Ser	
5700	5705	5710	
gtg aag ctg ccg ctg	acc gac cgc gag att	ccc atc atc gcg gac	80283
Val Lys Leu Pro Leu	Thr Asp Arg Glu Ile	Pro Ile Ile Ala Asp	
5715	5720	5725	
gcg gag ctg gtg gac	ccg aag ttc ggc acc	ggc gtg gtg aag gtg	80328
Ala Glu Leu Val Asp	Pro Lys Phe Gly Thr	Gly Val Val Lys Val	
5730	5735	5740	
acg ccg gcg cac gac	ttc aac gac tac cag	acg ggc ctg cgg cac	80373
Thr Pro Ala His Asp	Phe Asn Asp Tyr Gln	Thr Gly Leu Arg His	
5745	5750	5755	
aag ctg ccc atg ctc	acc atc ctg gac gaa	gcg gcc cgg atg acg	80418
Lys Leu Pro Met Leu	Thr Ile Leu Asp Glu	Ala Ala Arg Met Thr	
5760	5765	5770	
aag gag acc ggc aag	tac gcc ggc atg gat	cgc ttc gag gcg cgc	80463
Lys Glu Thr Gly Lys	Tyr Ala Gly Met Asp	Arg Phe Glu Ala Arg	
5775	5780	5785	
aag cag gtg ctg gcg	gac ctc acg gag cag	ggg ctg ctg gag aag	80508
Lys Gln Val Leu Ala	Asp Leu Thr Glu Gln	Gly Leu Leu Glu Lys	
5790	5795	5800	

gaa gag ccc cac atg	ctg aac gtg ggc acc	tgc cag cgc tgc gcc	80553
Glu Glu Pro His Met	Leu Asn Val Gly Thr	Cys Gln Arg Cys Ala	
5805	5810	5815	
acg gtg gtg gag cca	cgc ctg tct ccg cag	tgg ttc atc aag att	80598
Thr Val Val Glu Pro	Arg Leu Ser Pro Gln	Trp Phe Ile Lys Ile	
5820	5825	5830	
gaa ccg ctg gcg aag	ccg gcc atc gag gcg	gtg gag cag gcc cgc	80643
Glu Pro Leu Ala Lys	Pro Ala Ile Glu Ala	Val Glu Gln Gly Arg	
5835	5840	5845	
acg aag ttc gtc ccc	gag tcc tgg acg aac	acg ttc ttc cac tgg	80688
Thr Lys Phe Val Pro	Glu Ser Trp Thr Asn	Thr Phe Phe His Trp	
5850	5855	5860	
atg cgc aac atc cac	gac tgg tgc gtc agc	cgc cag ctg tgg tgg	80733
Met Arg Asn Ile His	Asp Trp Cys Val Ser	Arg Gln Leu Trp Trp	
5865	5870	5875	
ggc cac cag att ccc	gcg tac tac tgc acc	tcg tgc agc ccg cgg	80778
Gly His Gln Ile Pro	Ala Tyr Tyr Cys Thr	Ser Cys Ser Pro Arg	
5880	5885	5890	
cag ggc gac gac acg	gac ctg ccg ctg gac	gcc ccc acg gtg aag	80823
Gln Gly Asp Asp Thr	Asp Leu Pro Leu Asp	Ala Pro Thr Val Lys	
5895	5900	5905	
gtg ggc ggg gtg gac	ttc gcg cgc gcg gag	ccc atc gtc gcg cgc	80868
Val Gly Gly Val Asp	Phe Ala Arg Ala Glu	Pro Ile Val Ala Arg	
5910	5915	5920	
gag cag ccg aag tcc	tgc ccg aag tgc ggc	ggc gcg tcc ttc atc	80913
Glu Gln Pro Lys Ser	Cys Pro Lys Cys Gly	Gly Ala Ser Phe Ile	
5925	5930	5935	
cag gac ccg gac gtg	ctg gac acc tgg ttc	tcg tcc ggc ctg tgg	80958
Gln Asp Pro Asp Val	Leu Asp Thr Trp Phe	Ser Ser Gly Leu Trp	
5940	5945	5950	
ccg ttc tcc acg ctg	ggc tgg ccg cgt gag	acg ccg gag ctg aag	81003
Pro Phe Ser Thr Leu	Gly Trp Pro Arg Glu	Thr Pro Glu Leu Lys	
5955	5960	5965	
acc ttc tac ccg acg	tcc gtc atg gag acg	ggc cac gac atc atc	81048
Thr Phe Tyr Pro Thr	Ser Val Met Glu Thr	Gly His Asp Ile Ile	
5970	5975	5980	
ttc ttc tgg gtc gcc	cgg atg atg atg atg	ggc atc cac ttc atg	81093
Phe Phe Trp Val Ala	Arg Met Met Met Met	Gly Ile His Phe Met	
5985	5990	5995	
ggg gat gtg ccc ttc	cgc acc gtg tac ctg	cac gcg atg gtg cgc	81138
Gly Asp Val Pro Phe	Arg Thr Val Tyr Leu	His Ala Met Val Arg	
6000	6005	6010	
gac gag aag ggc cag	aag atg tcc aag acg	aag ggg aac gtc atc	81183
Asp Glu Lys Gly Gln	Lys Met Ser Lys Thr	Lys Gly Asn Val Ile	
6015	6020	6025	
gac ccc ctg gac atc	gtc ctc ggc gcc ccg	gcc gac aag ctc cag	81228
Asp Pro Leu Asp Ile	Val Leu Gly Ala Pro	Ala Asp Lys Leu Gln	

197

6030	6035	6040	
ccg acg ctg aag aac Pro Thr Leu Lys Asn 6045	cgg ttc ccg cag ggc Arg Phe Pro Gln Gly 6050	atg ccg ccg ttc ggc Met Pro Pro Phe Gly 6055	81273
gcg gac gcg ctg cgc Ala Asp Ala Leu Arg 6060	ttc acg ctc gcc tcg Phe Thr Leu Ala Ser 6065	ctc acc cag cag ggc Leu Thr Gln Gln Gly 6070	81318
cgc gac atc aag ctg Arg Asp Ile Lys Leu 6075	tcg atg gac cgg ctc Ser Met Asp Arg Leu 6080	gag ggc tac aag gcc Glu Gly Tyr Lys Ala 6085	81363
ttc ggc aac aag ctg Phe Gly Asn Lys Leu 6090	tgg aac gcc agc cgc Trp Asn Ala Ser Arg 6095	ttc gcc ctg atg aac Phe Ala Leu Met Asn 6100	81408
atg ggc gac ttc cgg Met Gly Asp Phe Arg 6105	atg gag gcc ggc ccg Met Glu Ala Gly Pro 6110	ctg gac cgg agc aca Leu Asp Arg Ser Thr 6115	81453
ctg acg ctg gcg gac Leu Thr Leu Ala Asp 6120	cgc tgg att ctc tcc Arg Trp Ile Leu Ser 6125	cgc ctc cag cgc gcc Arg Leu Gln Arg Ala 6130	81498
acc acg gag gcg cgc Thr Thr Glu Ala Arg 6135	gcg ctg ctg gag tcg Ala Leu Leu Glu Ser 6140	tac agc ttc ggc gag Tyr Ser Phe Gly Glu 6145	81543
gcc gcc tcc acg ctc Ala Ala Ser Thr Leu 6150	tac cag ttc ctc tgg Tyr Gln Phe Leu Trp 6155	gcg gag ttc tgc gac Ala Glu Phe Cys Asp 6160	81588
tgg tac atc gag ctg Trp Tyr Ile Glu Leu 6165	gcg aag ggc tcg ctg Ala Lys Gly Ser Leu 6170	tac ggc gag gac gcg Tyr Gly Glu Asp Ala 6175	81633
gcg gcg aag gac tcc Ala Ala Lys Asp Ser 6180	acg cgc cgg gtg ctc Thr Arg Arg Val Leu 6185	gtg tac tgc ctg gac Val Tyr Cys Leu Asp 6190	81678
cgc atc ctg cgg ctg Arg Ile Leu Arg Leu 6195	ctg cac ccg ttc atg Leu His Pro Phe Met 6200	ccg ttc atc acc gag Pro Phe Ile Thr Glu 6205	81723
gag atc tgg cag aag Glu Ile Trp Gln Lys 6210	ctg ccg atg gcc cgg Leu Pro Met Ala Arg 6215	ccg acg gac agc atc Pro Thr Asp Ser Ile 6220	81768
atg att gcc ccg tac Met Ile Ala Pro Tyr 6225	ccg gag ccg gac gcg Pro Glu Pro Asp Ala 6230	gcg cac gtg gac gag Ala His Val Asp Glu 6235	81813
gcg gcg gaa ggc gag Ala Ala Glu Gly Glu 6240	atg gcg ccg gtc atc Met Ala Pro Val Ile 6245	gcc gcc atc gag ggc Ala Ala Ile Glu Gly 6250	81858
ctg cgc acc atc cgc Leu Arg Thr Ile Arg 6255	ggc gag agc aac ctg Gly Glu Ser Asn Leu 6260	tcg ccc gcc acc aag Ser Pro Ala Thr Lys 6265	81903
gtg aag gcg gtg gtg	cag agc gcg gac gtg	cgc acg cgc gag ctg	81948

198

Val	Lys	Ala	Val	Val	Gln	Ser	Ala	Asp	Val	Arg	Thr	Arg	Glu	Leu	
				6270					6275					6280	
ctg	gag	cgc	tgg	cgc	ggc	tac	ctc	atg	ccg	ctg	gcc	ggc	ctg	tcc	81993
Leu	Glu	Arg	Trp	Arg	Gly	Tyr	Leu	Met	Pro	Leu	Ala	Gly	Leu	Ser	
				6285					6290					6295	
gag	gtg	acg	gtg	gga	gcg	ccg	ggc	gcc	aag	ccg	ccg	cag	gcc	gcc	82038
Glu	Val	Thr	Val	Gly	Ala	Pro	Gly	Ala	Lys	Pro	Pro	Gln	Ala	Ala	
				6300					6305					6310	
gcc	ttc	gtg	ggc	tcc	aac	ctg	gag	atc	tac	gtc	ccg	ctg	gcg	ggc	82083
Ala	Phe	Val	Gly	Ser	Asn	Leu	Glu	Ile	Tyr	Val	Pro	Leu	Ala	Gly	
				6315					6320					6325	
ctc	gtc	gac	ctg	gac	gcg	gag	cgc	gag	cgg	ttg	aag	aag	gag	att	82128
Leu	Val	Asp	Leu	Asp	Ala	Glu	Arg	Glu	Arg	Leu	Lys	Lys	Glu	Ile	
				6330					6335					6340	
gcc	cgc	gcc	gag	cag	gag	atg	gcc	ggc	gtg	ctg	cgc	aag	ctg	gag	82173
Ala	Arg	Ala	Glu	Gln	Glu	Met	Ala	Gly	Val	Leu	Arg	Lys	Leu	Glu	
				6345					6350					6355	
aac	ccc	aac	ttc	gtg	gcc	aag	gcg	ccc	ccg	gac	gtg	gtg	gag	aag	82218
Asn	Pro	Asn	Phe	Val	Ala	Lys	Ala	Pro	Pro	Asp	Val	Val	Glu	Lys	
				6360					6365					6370	
gac	cgc	gcc	cgc	gtc	gag	gag	ctg	aag	gag	cgc	aag	gcc	aag	ctt	82263
Asp	Arg	Ala	Arg	Val	Glu	Glu	Leu	Lys	Glu	Arg	Lys	Ala	Lys	Leu	
				6375					6380					6385	
cag	gac	cac	ctg	cag	cgg	att	gcc	ccg	gag	ccc	gcc	atg	cca	gag	82308
Gln	Asp	His	Leu	Gln	Arg	Ile	Ala	Pro	Glu	Pro	Ala	Met	Pro	Glu	
				6390					6395					6400	
acg	ccg	ccg	tcc	gag	agc	acc	act	ccc	gag	agc	agc	acc	ccc	acc	82353
Thr	Pro	Pro	Ser	Glu	Ser	Thr	Thr	Pro	Glu	Ser	Ser	Thr	Pro	Thr	
				6405					6410					6415	
gcg	gaa	gag	ggt	gcc	tcc	ccc	tcc	gcc	ccg	gcc	cag	gtg	aag	gtc	82398
Ala	Glu	Glu	Gly	Ala	Ser	Pro	Ser	Ala	Pro	Ala	Gln	Val	Lys	Val	
				6420					6425					6430	
gcc	ccg	aat	cag	gaa	gag	aag	ggc	ggc	gtg	gac	ctg	ggc	cag	gag	82443
Ala	Pro	Asn	Gln	Glu	Glu	Lys	Gly	Gly	Val	Asp	Leu	Gly	Gln	Glu	
				6435					6440					6445	
ctg	aag	ggt	gaa	ctg	ggc	gag	acc	ggc	acc	gcg	tcc	gag	gcc	ggg	82488
Leu	Lys	Gly	Glu	Leu	Gly	Glu	Thr	Gly	Thr	Ala	Ser	Glu	Ala	Gly	
				6450					6455					6460	
gac	acg	cag	gtg	cag	gac	gcg	ctc	gcg	aag	ctg	cgc	gag	ggc	acg	82533
Asp	Thr	Gln	Val	Gln	Asp	Ala	Leu	Ala	Lys	Leu	Arg	Glu	Gly	Thr	
				6465					6470					6475	
aag	gag	ggc	ctg	tcg	cca	tcc	gac	cac	cat	gac	ctc	ggc	gtg	gcg	82578
Lys	Glu	Gly	Leu	Ser	Pro	Ser	Asp	His	His	Asp	Leu	Gly	Val	Ala	
				6480					6485					6490	
tac	atg	acc	atg	ggg	ctg	gtg	gac	gac	gcg	atg	cgc	gag	ttc	aac	82623
Tyr	Met	Thr	Met	Gly	Leu	Val	Asp	Asp	Ala	Met	Arg	Glu	Phe	Asn	
				6495					6500					6505	

199

cg	g	a	g	g	g	g	a	a	c	c	g	g	a	c	c	g	g	a	g	g	c		82668
Arg	Ala	Lys	Glu	Gly	Gly	Asp	Thr	Arg	Glu	Thr	Pro	Ala	Glu	Gly									
				6510					6515					6520									
g	c	a	g	g	g	g	a	t	c	c	c	g	g	a	a	g	g	c	c	a	c	g	82713
Ala	Glu	Ala	Glu	Glu	Gly	Glu	Ser	Gly	Pro	Val	Lys	Ala	His	Gly									
				6525					6530					6535									
g	a	c	a	t	c	c	a	c	g	c	g	a	g	g	t	a	c	g	g	a	c	c	82758
Asp	Met	Leu	Pro	Thr	Arg	Glu	Val	Thr	Gln	Gly	Glu	Pro	Val	Arg									
				6540					6545					6550									
t	c	c	a	a	a	a	g	c	a	c	c	g	a	a	g	g	t	a	c	c	g	82803	
Ser	Ala	Ser	Lys	Lys	Lys	Ala	Thr	Ala	Pro	Gly	Lys	Val	His	Gly									
				6555					6560					6565									
g	a	a	c	t	g	g	c	c	c	a	c	a	c	g	a	g	g	c	g	a	g	c	82848
Asp	Lys	Leu	Ala	Pro	Ala	Pro	Thr	Thr	Glu	Glu	Ala	Glu	Ala	Pro									
				6570					6575					6580									
g	c	g	g	t	a	a	a	g	g	c	g	g	c	g								82868	
Ala	Val	Lys	Lys	Ala	Ala																		
				6585																			

<210> 19

<211> 358

<212> PRT

<213> Angiococcus disciformis

<400> 19

Val	Gly	Arg	Gly	Leu	Gln	Pro	Leu	Pro	Pro	Asp	Ala	Pro	His	Pro	Arg
1				5					10					15	

Pro	Ala	Arg	Cys	Glu	Gly	Arg	Arg	Trp	Gln	Gly	Leu	Gly	Ala	His	Gly
			20					25					30		

Gln	Pro	His	Leu	Arg	Pro	Arg	Arg	Gly	Glu	Ser	Gly	Pro	Gly	Arg	Cys
		35					40					45			

Arg	Asp	Gly	Ser	His	Arg	Glu	Gly	Thr	His	Pro	Pro	Gln	Ser	Arg	Ser
	50					55					60				

Tyr	Leu	Ala	Met	Thr	Ala	Asp	Ser	Met	Asp	Ala	Thr	Val	Pro	Ser	Arg
65					70				75						80

Arg	Pro	Ala	Leu	Gly	Phe	Ile	Gly	Gly	Gly	Val	Met	Ala	Glu	Ala	Leu
				85					90					95	

Ile	Gly	Ala	Val	Thr	Gly	Gly	Gly	Leu	Ala	Ala	Pro	Gly	Thr	Leu	Leu
			100					105					110		

Val Gly Glu Val Asp Glu Glu Arg Arg Arg Val Leu Arg Glu Gln Tyr
 115 120 125

Gly Val Glu Val Thr Ala Asp Asn Leu Ala Ala Ala Asp Ala Glu Leu
 130 135 140

Val Val Leu Ala Val Lys Pro Gln His Leu Gly Val Val Leu Gln Gln
 145 150 155 160

Leu Arg Gly Gln Leu Arg Glu Asp Thr Ile Val Val Ser Ile Ile Ala
 165 170 175

Gly Ala Arg Leu Gln Thr Leu Cys Glu Gly Leu Glu His Pro Leu Val
 180 185 190

Val Arg Ala Met Pro Asn Thr Pro Ala Arg Val Arg Lys Ala Ala Thr
 195 200 205

Phe Trp Ile Pro Ser Ala Gly Leu Ser Glu Pro Ala Leu Val Arg Val
 210 215 220

Arg Ser Leu Leu Gly Ala Leu Gly Thr Glu Val Glu Val Ala Asp Glu
 225 230 235 240

Ala Gly Val Glu Met Ala Thr Gly Leu Ala Gly Pro Met Pro Ala Phe
 245 250 255

Val Tyr Phe Leu Ile Glu Ala Phe Ile Asp Ala Gly Val Ala Leu Gly
 260 265 270

Leu Pro Arg Asp Gln Ala Thr Leu Ala Thr Val Glu Ser Met Arg Gly
 275 280 285

Ser Leu Glu Leu Leu His Arg Ser Gln Glu Ser Pro Ala Ala Leu Arg
 290 295 300

Lys Gln Val Thr Ser Pro Gly Gly Ala Thr Leu Ala Gly Leu Lys Val
 305 310 315 320

Phe Glu Ser Ala Gly Val Arg Asp Ala Met Asp Ala Ala Val Arg Ala
 325 330 335

Val His Ala Arg Ala Gln Glu Leu Gly Arg Leu Pro Leu Ala Pro Leu
 340 345 350

Glu Pro Gly Lys Glu Gly

355

<210> 20

<211> 420

<212> PRT

<213> Angiococcus disciformis

<400> 20

Val Pro Gly Val Ala Val Ala Val Val Ser His Gly Arg Val Ile Val
1 5 10 15

Ala Lys Gly Tyr Gly Phe Ser Asn Leu Glu His Gln Val Pro Val Thr
20 25 30

Pro Asp Thr Leu Phe Gln Ser Gly Ser Leu Gly Lys Gln Phe Thr Ala
35 40 45

Met Ala Val Met Leu Gln Val Glu Ala Gly Arg Ile Ser Leu Ser Asp
50 55 60

Ser Ile Thr Lys Tyr Phe Pro Asp Ala Pro Ser Thr Trp Ala Pro Ile
65 70 75 80

Thr Ile Arg His Leu Leu Thr His Thr Ser Gly Leu Ser Asp Gly Leu
85 90 95

Leu Asp Leu Arg Lys Asp Tyr Thr Asp Glu Glu Phe Ala Gln Phe Leu
100 105 110

Tyr Thr Leu Pro Leu Asp Phe Pro Ala Gly Leu Arg Trp Asn Tyr Ser
115 120 125

Asn Ala Gly Tyr Val Leu Leu Gly Leu Leu Val Asn Arg Val Ala Gly
130 135 140

Thr Ser Tyr Val Asn Val Leu Gly Glu Gln Val Phe Lys Pro Thr Arg
145 150 155 160

Met Lys Thr Ala Arg Gly Ile Ser Glu Ala Asp Ile Val Pro Asn Arg
165 170 175

Ala Ala Gly Tyr Arg Leu Leu Asp Gly Val Val Lys Asn Gln Glu Trp
180 185 190

202

Val Ser Pro Ser Leu Asn Thr Thr Ala Asp Gly Ser Leu Tyr Phe Ser
195 200 205

Leu Lys Asp Met Leu Ala Trp Asp Asp Ala Val Asp Asp Arg Ala Ile
210 215 220

Leu Thr Asn Gly Ser Trp Arg Asp Ile Leu Ser Pro Val Lys Leu Ser
225 230 235 240

Ser Gly Ala Ser Tyr Pro Tyr Gly Phe Gly Trp Phe Ile Glu Glu Arg
245 250 255

Asn Gly Gln Leu Val His Gln His Ala Gly Ala Trp Gln Gly Phe Lys
260 265 270

Thr Ala Tyr Tyr Arg Phe Thr Gly Asp Ala Leu Ser Ile Ile Val Leu
275 280 285

Leu Asn Leu Ala Glu Pro Asn Pro Ala Ala Ile Ala Asn Gly Ile Ala
290 295 300

Ala Ile Val Asn Pro Ala Leu Ala Val Pro Pro Leu Ala Pro Ile Pro
305 310 315 320

Asp Leu Glu Pro Glu Val Thr Ala Arg Leu Ala Thr Leu Leu Glu Gln
325 330 335

Ala Arg Ala Gly Ala Leu Asn Pro Ala Glu Phe Ala Tyr Val Pro Gly
340 345 350

Trp Phe Phe Thr Glu Ala Ala Pro Tyr Tyr Gln Ser Leu Leu Gln Ser
355 360 365

Leu Gly Pro Ser Gly Pro Leu Val Leu Ala Lys Arg Glu Met Arg Gly
370 375 380

Asp Asp Arg Val Tyr Thr Tyr Leu Val Gln Ala Gly Ser Thr Met Leu
385 390 395 400

Arg Tyr Arg Val Ala Leu Ile Pro Asp Gly Arg Val Ser Ala Phe Ala
405 410 415

Leu Ser Pro Asn
420

<210> 21

<211> 184

<212> PRT

<213> *Angiococcus disciformis*

<400> 21

Val Ser Ile Arg Arg Thr Thr Ala Ser Thr Arg Asn Val Thr Val Ser
1 5 10 15

Ala Met Leu Asp Pro Ala Leu Gly Ser Ala Arg Asp Ala Leu Gln Val
20 25 30

Phe Leu Glu Lys Ala Leu Arg Ala Gln Phe Gly Ala Asp Trp Gln Gly
35 40 45

Thr Leu Glu Arg Asp Arg Leu Ala Gln Gly Arg Pro Thr Ser Glu Arg
50 55 60

Trp Asp Trp Gly Ser Ile Ser Tyr Arg Leu Gly His Phe Arg Ser Pro
65 70 75 80

Gln Ala Thr Gly Ser Gly Val Val Thr Glu Pro Ile Val Ala Leu Ala
85 90 95

Arg Leu Ala Lys Ala Tyr Arg Asn Leu Phe Gln His Glu Glu Leu Thr
100 105 110

Asp Ala Gln Val Arg His Ala Met Glu Gly Leu Ala Leu Leu Gln Glu
115 120 125

Ala Leu Gly Asn Lys Thr Arg Ala Ala Glu Thr Arg Ala Ala Leu Glu
130 135 140

Val Trp Ala Gly Pro Pro Ala Arg Ala Ala Lys Arg Pro Ala Arg Lys
145 150 155 160

Ala Pro Ala Ala Ala Lys Arg Ala Lys Pro Ala Ala Lys Lys Lys Ala
165 170 175

Pro Lys Ala Lys Ser Ala Arg Arg
180

<210> 22

<211> 351

<212> PRT

<213> Angiococcus disciformis

<400> 22

Met Thr Ser Gly Thr Arg Lys Thr Gly Gly Asp Gly Ala Arg Arg Asn
1 5 10 15

Asp Lys Tyr Arg Ile Leu Ser Leu Asp Gly Gly Gly Leu Arg Thr Leu
20 25 30

Leu Ser Ile Gly Leu Leu Lys Arg Ile Asp Gln Leu Arg Pro Gly Phe
35 40 45

Leu Asp Gln Val Asp Leu Val Ala Gly Thr Ser Ala Gly Ala Ile Ser
50 55 60

Ala Leu Ile Ile Ala Ala Ala Arg Glu Pro Ala Val Gly Leu Glu Gln
65 70 75 80

Ala Arg Gln Ile Trp Phe Thr Pro Gly Leu Phe Asp Ser Pro Leu Ser
85 90 95

Asn Gln Leu Gly Ala Leu Val Gly Gln Ser Ala Leu Met Pro Ser Glu
100 105 110

Asn Met Ala Arg Ala Leu Thr His Ile Leu Gly Asp Lys Thr Leu Arg
115 120 125

Asp Leu Lys Arg Lys Val Val Ile Pro Ala Phe Gln Leu Asp Asp Gly
130 135 140

Asp Pro Asp Glu Asp Arg Arg Gly Trp Arg Pro Arg Ile Phe His Asn
145 150 155 160

Phe Pro Gly Asp Thr Phe Val Asn Leu Asp Asp Tyr Leu Val Asp Leu
165 170 175

Ala Leu Arg Ser Ser Ser Leu Pro Ile Val Ser Pro Val His Gln Gly
180 185 190

Tyr Val Asp Gly Gly Leu Phe Ala Asn Asn Pro Thr Met Ser Ala Val
195 200 205

Ala Gln Ala Ile Tyr Ala Lys Ala Ala Asp Val Arg Asp Ile Leu Val
210 215 220

Phe Ser Leu Gly Thr Gly Asp Ser Val Asp Tyr Leu Asp Gly Tyr Asn
225 230 235 240

Glu Asn Trp Gly Trp Arg Lys Trp Leu Leu Asp Pro Lys Gln Pro Met
245 250 255

Ala Phe Val Ala Ala Thr Ile Glu Ala Gly Val Glu Ala Ile Asp Phe
260 265 270

Gln Ala Lys Met Leu Leu Pro Arg Gly Asn Tyr Trp Arg Glu Asp Pro
275 280 285

Val Val Pro Phe Asn Leu Gly Asn Ser Val Glu Ala Gln Ile Ala Thr
290 295 300

Leu Asp Arg Val Val Lys Arg His Asp Leu Ser Lys Val Leu Glu Trp
305 310 315 320

Val Glu Asn Ser Gly Trp Gln Pro Lys Ala Thr Gln Glu Thr Ala Ser
325 330 335

Gly Ala Gln Arg Glu Phe Ser Gly Glu Ser Gly Glu Ala Arg Ser
340 345 350

<210> 23

<211> 351

<212> PRT

<213> *Angiococcus disciformis*

<400> 23

Met Thr Ile Ile His Gly Gln Pro Gly Ala Asp Gly Ala Arg Arg Asn
1 5 10 15

Gly Lys Tyr Arg Ile Leu Ser Leu Asp Gly Gly Gly Leu Arg Thr Leu
20 25 30

Leu Ser Ile Gly Leu Leu Lys Arg Ile Asp Gln Leu Arg Pro Gly Phe
35 40 45

Leu Asp Gln Val Asp Leu Val Ala Gly Thr Ser Ala Gly Ala Ile Ser
50 55 60

Ala Leu Ile Ile Ala Ala Ala Arg Glu Pro Ala Val Gly Leu Glu Gln
65 70 75 80

Ala Arg Gln Leu Trp Leu Thr His Glu Leu Phe His Ser Ser Leu Arg
85 90 95

Thr Arg Leu Gly Ala Ile Ile Gly Met Asn Ala Leu Thr Pro Ser Lys
100 105 110

Asn Met Ala Arg Ala Leu Thr His Ile Leu Gly Asp Lys Thr Leu Gly
115 120 125

Asp Leu Lys Arg Lys Val Val Ile Pro Ala Phe Lys Leu Asp Asp Glu
130 135 140

Asn Pro Asp Glu Asn Arg Arg Glu Trp Lys Pro Arg Val Phe His Asn
145 150 155 160

Phe Pro Gly Ala Asp Phe Val Asn Pro Gly Asp Arg Leu Val Asp Leu
165 170 175

Ala Leu Arg Ser Ser Ser Leu Pro Ile Val Ser Pro Val Tyr Gln Gly
180 185 190

Tyr Val Asp Gly Gly Leu Phe Ala Asn Asn Pro Thr Met Ala Ala Val
195 200 205

Ala Gln Ala Ile Tyr Ala Lys Ala Ala Asp Leu Arg Asp Leu Leu Val
210 215 220

Phe Ser Gln Gly Ala Gly Asp Met Val His Tyr Leu Lys Gly Tyr Asn
225 230 235 240

Glu Asn Trp Gly Trp Trp Lys Trp Leu Leu Asn Ser Lys Leu Pro Trp
245 250 255

Ala Leu Val Asp Ile Ala Val Ala Ala Gly Val Glu Ala Ile Asp Phe
260 265 270

Gln Ala Lys Arg Leu Leu Pro Arg Gly Asn Tyr Trp Arg Glu Asp Pro
275 280 285

Arg Val Pro Ser His Leu Gly Asn Ser Val Glu Ala Gln Val Val Thr
290 295 300

Leu Asp Gln Ile Val Glu Lys His Asp Leu Ser Glu Ala Leu Glu Trp
305 310 315 320

Val Asp Ser Ser Gly Trp Leu Pro Lys Gly Met His Glu Ala Ala Ser

325

207
330

335

Gly Ala Gln Pro Glu Leu Ser Gly Glu Ser Gly Glu Ala Arg Ser
340 345 350

<210> 24

<211> 405

<212> PRT

<213> *Angiococcus disciformis*

<400> 24

Met Ala Ala Arg Cys Arg Gln Ala Arg Gly Ala Ser Met Thr Arg Ile
1 5 10 15

Leu Leu Phe Ser Gly Lys Gly Gly Val Gly Lys Thr Thr Ala Ser Ala
20 25 30

Ala Thr Ala Val Ala Ala Ala Lys Arg Gly Tyr Arg Thr Leu Val Met
35 40 45

Ser Phe Asp Ile Ala His Ser Leu Ser Asp Ser Phe Asp Leu Asp Arg
50 55 60

Lys Leu Phe Asp Phe Asn Glu Gly Leu Pro Gln Lys Val Ala Pro Asn
65 70 75 80

Leu Glu Leu Gln Glu Ile Asp Ile Gln His Glu Leu Gln Arg Gln Trp
85 90 95

Ser Glu Val Tyr Asn Tyr Met Ser Val Leu Leu Thr Ser Thr Gly Val
100 105 110

Ser Asn Met Val Ala Glu Glu Val Ala Ile Leu Pro Gly Thr Glu Asp
115 120 125

Val Ile Ser Leu Ile Tyr Leu Asn Gln Tyr Val Gln Gln Gly Arg Tyr
130 135 140

Asp Val Ile Ile Val Asp Cys Pro Pro Thr Gly Glu Ser Leu Arg Phe
145 150 155 160

Val Asn Ile Thr Ser Val Leu Gln Trp Tyr Ile Arg Arg Arg Phe Asn
165 170 175

208

Val Asp Arg Thr Leu Val Lys Leu Ala Arg Pro Leu Ala Ser Arg Leu
180 185 190

Thr Ser Tyr Asp Leu Pro Glu Asp Ser Tyr Phe Ala Ser Leu Glu Gln
195 200 205

Leu Phe Asp Arg Val Lys Gly Ile Glu Gly Leu Leu Thr Asp Gly Asn
210 215 220

Arg Thr Thr Val Arg Leu Val Ser Ser Ala Glu Lys Met Val Ile Arg
225 230 235 240

Glu Thr Gln Arg Ala Tyr Leu Tyr Phe Asn Met Tyr Gly Met Thr Val
245 250 255

Asp Gln Val Val Val Asn Arg Leu Leu Pro Asp Thr Glu Tyr Phe Ala
260 265 270

Gln Trp His Lys Ser Gln Ser Ala Tyr Leu Asp His Ile Gln Glu Tyr
275 280 285

Phe Ser Pro Met Pro Val Ala Arg Leu Pro Leu Leu Glu His Glu Val
290 295 300

Val Gly Leu Glu Arg Leu Glu Asp Leu Ala Gln Arg Leu Tyr Gly Asp
305 310 315 320

Ser Asp Pro Thr Glu Arg Tyr Ile Ser Gly Gln Pro Tyr Gln Phe Ala
325 330 335

Lys Gln His Glu Gly Arg Tyr Arg Leu Gln Leu Leu Met Pro Gly Ala
340 345 350

Asp Arg Gly Asp Ile Val Leu Asp Arg Gln Gly Asn Asp Leu Ile Ile
355 360 365

Arg Val Gly Gly Phe Arg Arg His Val Met Leu Pro Arg Ser Val Ala
370 375 380

Gln Leu Asp Ala Val Asp Ala Val Leu Asp Lys Gly Met Leu Lys Ile
385 390 395 400

Asp Phe Ala Arg Pro
405

<210> 25

<211> 373

<212> PRT

<213> Angiococcus disciformis

<400> 25

Val Asn Ala Val Arg Pro Val Met Gln Leu Lys Gly Ala Gly Ser Ala
1 5 10 15

Met Arg Met Thr Arg Ile Leu Gly Glu Gln Glu Leu Arg Gln Ile Val
20 25 30

Arg Ala Thr Gly Leu His Ala Leu Met Asp Glu Val Ile Ser Thr Leu
35 40 45

Glu Tyr Glu Leu Glu His Phe Asp Pro Ala Asn Thr Glu Leu Arg Lys
50 55 60

Arg Glu Gly Phe Gln Tyr Gln His Pro His Pro Gly Val Leu Glu Trp
65 70 75 80

Met Pro Val Met Gln Val Gly Arg His Ala Val Ile Lys Leu Val Gly
85 90 95

Tyr Asn Pro Arg Ser Pro Glu Glu Leu Gly Leu Pro Thr Ile Leu Ser
100 105 110

Ser Leu Ser Leu Tyr Asp Val Thr Thr Gly His Leu Arg Ala Val Cys
115 120 125

Asp Gly Thr Phe Pro Thr Ala Leu Arg Thr Gly Ala Ala Ser Ala Val
130 135 140

Ala Ser Arg Leu Leu Ala Arg Glu Asp Ser Arg Val Leu Gly Leu Val
145 150 155 160

Gly Cys Gly Ala Gln Ala Val Thr Gln Ala His Ala Leu Ser Arg Val
165 170 175

Phe Pro Leu Gly Arg Val Leu Val His Asp Ile Ala Pro Glu Ala Met
180 185 190

Arg Ser Phe Ala Arg Arg Val Ala Phe Leu Gly Leu Glu Val Val Ala
195 200 205

Val Pro Leu Thr Arg Val Glu Gln Glu Ala Asp Ile Leu Cys Thr Ala

210

210

215

220

Thr Ser Val Ala Val Gly Ala Gly Pro Val Val Ser Gly Arg Leu Ile
 225 230 235 240

Lys Ala His Leu His Ile Asn Ala Val Gly Ser Asp Leu Pro Gly Lys
 245 250 255

Thr Glu Leu Pro Leu Glu Met Leu Arg Arg Ala Leu Val Cys Pro Asp
 260 265 270

Tyr Leu Pro Gln Ala Val Val Glu Gly Glu Cys Gln Gln Leu Glu Pro
 275 280 285

Gly Glu Ile Gly Pro Ser Leu Phe Glu Leu Thr Arg Ser Pro Glu Gln
 290 295 300

Trp Arg Ser His Arg Asp Gly Leu Thr Val Phe Asp Ser Thr Gly Phe
 305 310 315 320

Ser Leu Glu Asp Lys Val Val Leu Glu Val Ile Leu Ala His Ala Glu
 325 330 335

Arg Leu Gly Val Gly Arg Leu Leu Gln Val Glu Ser Met Ala Glu Asp
 340 345 350

Pro Lys Asp Pro Tyr Gly Phe Leu Gln Glu Thr Ala Thr Cys Leu Lys
 355 360 365

Ala Ala Pro Ala Val
 370

<210> 26

<211> 219

<212> PRT

<213> *Angiococcus disciformis*

<400> 26

Met Arg Ile Leu Thr Asp Val Phe Leu Arg Lys Tyr Gln Arg Ala Pro
 1 5 10 15

Gly Pro Arg Glu Arg Ser Leu Ala Ala Asp Ala Arg Ser Lys Leu Trp
 20 25 30

211

Leu Ser Val Gly Leu Thr Ala Phe Leu Val Pro Ala Tyr Ile Glu Tyr
 35 40 45

Phe Cys Arg Met Leu Gly Leu Pro Lys Gly Glu Leu Pro Trp Lys Leu
 50 55 60

Tyr Asp Met Leu Arg Glu Arg Asn Glu Ala Pro Asp Leu Lys Gln Arg
 65 70 75 80

Phe Gln Gln Leu Glu His Leu Pro Glu Gly Ser Phe Gly Arg Glu Phe
 85 90 95

Trp Lys Tyr Cys His Ser His Gly Val Pro Leu Pro Gly His Pro His
 100 105 110

Ala Leu Pro Val His Ser Ala Val His Asp Phe Val His Leu Leu Ser
 115 120 125

Gly Tyr Gly Ile Ser Ile Trp Glu Glu Met Leu Thr Thr Ala Phe Ser
 130 135 140

Leu Gly Phe Met Phe Gly Pro Arg Thr Arg Asn Tyr Asp Pro Ala Gln
 145 150 155 160

Tyr Lys Ser Leu Val Pro Val Arg Leu Leu Ala Glu Ala Leu Pro Ile
 165 170 175

Pro Gln Leu Arg Arg Ala Val Glu Arg Gly Ser Ala Met Ser Val Asp
 180 185 190

Leu Phe Gly Asp Trp Asp Pro Trp Lys Val Met Glu Leu Pro Leu Glu
 195 200 205

Glu Val Lys Arg Met Tyr Asn Ile Gln Pro Glu
 210 215

<210> 27

<211> 781

<212> PRT

<213> Angiococcus disciformis

<400> 27

Met Gly His Glu Asp Ser Asp Lys Ala Leu Pro Asp Ser Lys Ala Asn
 1 5 10 15

Glu Val Glu Ser Ser Arg Asn Glu Ala Glu Trp Asn Glu Leu Pro Ile
 20 25 30

Ser Ala Arg Asp Cys Glu Asp Leu Lys Lys Ser Cys Leu Leu Pro Ser
 35 40 45

Thr Val Arg Lys Leu Gly Trp Arg Thr Asp Thr Asp Glu Asn Asn Arg
 50 55 60

Lys Leu Leu Asn Trp Pro Glu Glu Trp Lys Gly Asp Ala Gly Gly Leu
 65 70 75 80

Val Leu Val Tyr Pro Gly Leu Asn Tyr Ala Lys Val Lys Pro His Arg
 85 90 95

Pro Arg Val Leu Arg Pro Lys Glu Ala Gln Leu Thr Leu Pro Thr Ala
 100 105 110

Thr Phe Thr Asp His Ile Leu Arg Pro Tyr Thr Gln Asp Glu Val Glu
 115 120 125

Gly Glu Pro Ala Lys Val Lys Lys Tyr Glu Ala Pro Phe Gly Tyr Ala
 130 135 140

Pro Gln Ile Phe Asp Thr Ala Ile Leu Cys Pro Met Leu Gln Asp Glu
 145 150 155 160

Glu Val Pro Met Ala Leu Ile Glu Gly Glu Lys Lys Ala Ala Leu Leu
 165 170 175

Ala Gln Glu Gly Thr Pro Ala Leu Gly Leu Pro Gly Ile Ser Met Ala
 180 185 190

His Asp Val Glu Ala Arg Arg Arg Ser Trp Asp Asp Leu Thr Asp Glu
 195 200 205

Trp Val Pro His Pro Glu Ala Ala Ser Tyr Leu Ala Gln Gly Arg Thr
 210 215 220

Val Phe Val Cys Gly Asp Ala Pro Asp Met Asp Gly Thr Asn Met Asn
 225 230 235 240

Val Val Arg Glu Val Val Arg Ile Ala Lys Gly Leu Ser Ala Arg Gly
 245 250 255

Cys Asp Pro Gln Ile Pro Tyr Val Asp His Asp Pro Gln His Glu Lys
 260 265 270

Ser Gly Val Asp Asp Trp Val His Arg Glu Met Lys Lys Leu Leu Ala
 275 280 285

Ala His His His Lys Thr Glu Ala Trp Ala Arg Gln Phe Leu Ile Val
 290 295 300

Gln Tyr Leu Ala Ser Ala Arg Pro Ala Leu Pro Ser Ala Leu Cys Asp
 305 310 315 320

Trp Leu Val Glu Gln His Glu Ala His Gly Gly Val Lys Glu Ile Glu
 325 330 335

Ile Lys Arg Ala Leu Val Trp Cys Ser Val Trp Phe Ser Thr Glu His
 340 345 350

Gly Ile Phe Lys Lys Trp Val Gly Thr Leu Lys Arg Lys Lys Val Cys
 355 360 365

Lys Pro Glu Val Ile Asp Arg Leu Val Gln Ala Leu Pro Arg Val Glu
 370 375 380

Asn Arg Phe Glu Pro Arg Thr Trp Val Gln His Trp Met Glu Lys Asn
 385 390 395 400

Gly Val Thr Phe Val Tyr Ala Ser Asp Gly Phe Gln Ile Gly Gly Arg
 405 410 415

Met Phe Pro Ala Ser Glu Phe Leu Ser Gln Ile Glu Ile Glu Ser Thr
 420 425 430

Asn Glu Arg Thr Cys Ile Ser Val Arg Asn Leu Gly Asn Ala Val Arg
 435 440 445

Leu Trp Glu Ala Asp Gln Arg Arg Gly Ile Val Glu Gln Tyr Arg Glu
 450 455 460

Lys Leu Lys Phe Arg Glu Asp Gly Arg Arg Gly Asp Ala Glu Glu Phe
 465 470 475 480

Ile Arg Ala Thr Thr Gly Arg Leu Asp Pro Val Asp Leu Val Val Val
 485 490 495

Leu His Phe Ile Trp Gln Val Lys Arg Lys Leu Phe Gly Lys Pro Val
 500 505 510

Lys Asp His Leu Met Pro Ile Phe Val Gly Lys Gln Arg Gly Gly Lys

515

520

525

Thr Thr Ala Ile Lys Ser Leu Leu Gln Pro Leu Gly Asp Leu Val Gly
530 535 540

Phe Pro Ala Asp Leu Gln Phe Leu Thr Asp Glu Arg Gln Gln Phe Arg
545 550 555 560

Leu Thr Arg Cys Tyr Gly Met Phe Phe Asp Glu Met Ala Lys Ala Glu
565 570 575

Arg Val Ala Ser Asp Val Leu Lys Asn Arg Ile Thr Ala Glu Arg Ile
580 585 590

Glu Tyr Arg Val Leu Gly Val Gln Lys Leu Gln Thr Gly Thr Asn Asn
595 600 605

Ala Thr Phe Ile Gly Ala Ser Asn Lys Gln Val Lys Glu Ile Val Val
610 615 620

Asp Pro Thr Gly Met Arg Arg Phe Tyr Gln Leu Asp Cys Leu Pro Gln
625 630 635 640

Met Asn Trp Thr Ala Leu Glu Gln Ile Asp Phe Leu Ala Met Trp Thr
645 650 655

Ser Val Asp Glu Lys Ala Glu Ser Pro Leu Leu Gly Gln Leu Ala Gln
660 665 670

Val Gln Ala Arg Gln Asp Glu Ile Arg Ala Lys Asp Ser Val Glu Asp
675 680 685

Phe Phe Glu Arg Arg Cys Ala Gln Arg Asp Gln Trp Ile Lys Ala Thr
690 695 700

Ala Leu Tyr Gly His Tyr Val Glu Asp Leu Lys Tyr Gln Gly Arg Asp
705 710 715 720

Arg Met Ala Phe Ser Leu Thr Lys Phe Gly Glu Arg Met Lys Glu Leu
725 730 735

Ala Gly Glu Glu Asn Thr Gly Trp Lys Arg Ser Asp Gly Ile Lys Tyr
740 745 750

Arg Val Thr Val Lys Arg Trp Glu Thr Glu Ile Asp Gly Val Met Glu
755 760 765

215

Ser Ile Ala Ala Gly Ala Arg Met Asp Leu Ala His Pro
 770 775 780

<210> 28

<211> 332

<212> PRT

<213> *Angiococcus disciformis*

<400> 28

Met Asp Thr Asp Lys Val Ala Leu Phe Gly Ala Ser Gly Pro Ile Gly
 1 5 10 15

His Ser Ile Val Lys Ala Leu Arg Glu Gln Gly Arg Ala Tyr Arg Val
 20 25 30

Val Gly Arg Ser Arg Ala Ser Leu Gln Ser Gln Phe Gly Ala Asp Pro
 35 40 45

Lys Ala Glu Ile Ala Thr Trp Asn Pro Asp Ala Pro Asp Thr Ile Arg
 50 55 60

Ala Ala Ala Arg Gly Ile Arg Thr Leu Ile Tyr Met Val Gly Val Asp
 65 70 75 80

Tyr Trp Arg Phe Gln Leu His Pro Glu Leu Met Arg Arg Thr Val Asp
 85 90 95

Ala Ala Ile Ala Glu Gly Val Glu Arg Val Val Leu Ile Gly Thr Val
 100 105 110

Tyr Pro Tyr Gly Arg Pro Arg Thr Thr Pro Val Thr Glu Asp His Pro
 115 120 125

Arg Glu Pro His Thr Phe Lys Gly Arg Met Arg Lys Glu Gln Glu Asp
 130 135 140

Ile Leu Leu Ala Ala Asp Ala Ala Gly Lys Leu Arg Ala Thr Ile Leu
 145 150 155 160

Arg Leu Pro Asp Phe Tyr Gly Pro Gly Ile Asp Lys Ser Phe Leu His
 165 170 175

Ser Ala Phe Val Ala Ala Ala Gln Gly Lys Arg Ala Gln Leu Ile Gly
 180 185 190

Pro Ile Asp Pro Pro His Glu Tyr Val Tyr Val Pro Asp Val Gly Pro
 195 200 205

Thr Val Thr Ala Leu Met Asp Ala Pro Gly Ala Tyr Gly Arg Trp Trp
 210 215 220

Asn Leu Ala Gly Ala Gly Val Thr Ser Gln Arg Lys Leu Val Glu Glu
 225 230 235 240

Ile Tyr Ala Gln Ala Gly His Leu Ala Lys Phe Met Thr Ala Gly Lys
 245 250 255

Gly Met Leu Arg Leu Ile Gly Leu Phe Asn Pro Phe Met Arg Glu Met
 260 265 270

Val Glu Met His Tyr Leu Leu Thr Glu Pro Val Leu Met Asp Asp Ser
 275 280 285

Ala Leu His Lys Leu Leu Gly Ser Val Arg Lys Thr Pro Tyr Ser Glu
 290 295 300

Gly Ile Arg Gln Thr Leu Glu Ala Thr Arg Ala Ser Leu Val Ser Pro
 305 310 315 320

Ala Pro Ala Ala Thr Gly Arg Pro His Pro Ala Pro
 325 330

<210> 29

<211> 547

<212> PRT

<213> Angiococcus disciformis

<400> 29

Met Ala Ala Lys Glu Ile Phe Phe His Gln Ser Ala Arg Glu Ala Ile
 1 5 10 15

Leu Arg Gly Val Arg Thr Leu Ala Asp Ala Val Ala Val Thr Leu Gly
 20 25 30

Pro Lys Gly Arg Asn Val Val Ile Glu Lys Ser Phe Gly Ser Pro Thr
 35 40 45

Ile Thr Lys Asp Gly Val Thr Val Ala Lys Glu Ile Asp Leu Asp Asn
 50 55 60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
 65 70 75 80

Thr Ser Asp Lys Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
 85 90 95

Arg Ala Ile Tyr Glu Glu Gly Leu Lys Leu Val Ala Ala Gly His Ser
 100 105 110

Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Glu Val Val Val
 115 120 125

Glu Glu Leu Lys Lys Leu Ser Lys Pro Thr Ala Asp Lys Lys Ala Ile
 130 135 140

Thr Gln Val Gly Thr Ile Ser Ala Asn Gly Asp Glu Thr Ile Gly Ser
 145 150 155 160

Ile Ile Ala Asp Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr
 165 170 175

Val Glu Glu Ala Lys Gly Leu Glu Thr Asn Leu Asp Val Val Glu Gly
 180 185 190

Met Gln Phe Asp Arg Gly Tyr Val Ser Pro Tyr Phe Val Thr Asn Arg
 195 200 205

Glu Arg Met Glu Val Val Met Asp Asp Pro Phe Ile Leu Ile Ser Glu
 210 215 220

Lys Lys Val Ser Ser Met Gln Asp Met Ile Pro Ile Leu Glu Gln Val
 225 230 235 240

Ala Arg Ser Gly Lys Pro Leu Leu Ile Ile Ala Asp Asp Ile Glu Gly
 245 250 255

Glu Ala Leu Ala Thr Leu Val Val Asn Lys Ile Arg Gly Val Leu Asn
 260 265 270

Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Glu Met
 275 280 285

Leu Lys Asp Ile Ala Thr Leu Thr Gly Gly Met Val Val Ser Glu Glu
 290 295 300

Leu Gly His Lys Tyr Glu Asn Leu Thr Leu Asn Asp Leu Gly Arg Ser

305

310

315

320

Lys Arg Ile Thr Val Asp Lys Asp Asn Thr Thr Ile Val Asp Gly Ala
325 330 335

Gly Thr Lys Ser Glu Ile Glu Gly Arg Ile Lys Leu Ile Arg Ser Gln
340 345 350

Ile Asp Thr Val Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg
355 360 365

Leu Ala Lys Leu Val Gly Gly Val Ala Val Ile Asn Val Gly Ala Ala
370 375 380

Thr Glu Thr Glu Met Lys Glu Lys Lys Ala Arg Val Glu Asp Ala Leu
385 390 395 400

His Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly
405 410 415

Val Ala Phe Leu Arg Cys Leu Pro Ala Leu Glu Lys Leu Lys Leu Gly
420 425 430

Gly Glu Gln Asp Phe Gly Val Glu Ile Ile Arg Arg Ala Leu Gln Glu
435 440 445

Pro Leu Arg Lys Ile Ala Ser Asn Ala Gly Val Glu Gly Ala Val Val
450 455 460

Ile Asn Lys Val Arg Glu Gly Gln Gly Ala Phe Gly Tyr Asn Ala Arg
465 470 475 480

Thr Glu Val Tyr Glu Asp Leu Glu Lys Ala Gly Val Ile Asp Pro Thr
485 490 495

Lys Val Glu Arg Thr Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu
500 505 510

Leu Leu Thr Thr Glu Ala Met Ile Ala Asp Arg Pro Ala Lys Lys Lys
515 520 525

Gly Lys Asn Gly Gly Gly Ala Gly Met Pro Asp Tyr Gly Gly Asp Asp
530 535 540

Met Asp Tyr
545

<210> 30

<211> 746

<212> PRT

<213> Angiococcus disciformis

<400> 30

Val Thr Arg Arg Cys Arg Cys Pro Ser Cys Pro Pro Arg Ala Gly Gly
1 5 10 15

Ala Cys Ser Val Ala Trp Pro Ala Pro Tyr Arg Ala Arg Pro Arg Ala
20 25 30

Pro His Ser Phe Arg His Pro Pro Cys Pro Arg Arg Pro Leu Pro His
35 40 45

Pro Cys Pro His Arg Ser Arg Ser His His Pro Ala Pro Cys Leu Arg
50 55 60

Pro Pro Ser Gly Pro Arg Pro Leu Arg Pro Ser Arg Ser Pro Pro Ser
65 70 75 80

Arg Pro Arg Pro Ala Ala Pro Ser Ala Gly Ala Cys Pro Arg Asn Arg
85 90 95

Pro Pro Arg Arg Ala Gly Arg Thr Gly Ala Arg Val Ser Trp Asp Leu
100 105 110

Ser Leu Glu Gly Ala Ala Gln Leu Thr Ala Glu Arg Arg Gly Ile Gly
115 120 125

Gln Gln Leu Arg Lys Ala Arg Gly Gly Thr Gly Ala Gln Arg Leu Gln
130 135 140

Arg Arg Leu His Gly Glu Val Leu Gly Gln His Asp Glu Gly Gly Ala
145 150 155 160

Leu Gly Gln Arg Leu Gly Gln Leu Gln Arg Leu Ala Pro Ile Gly Arg
165 170 175

Arg Glu Val Pro Ala His His His Ala His Arg Arg Arg His Arg His
180 185 190

Gln Leu Arg Arg Gly Gln His Pro Arg Leu Gln Pro Gln His Val Gln
195 200 205

Gly Leu Gly Asp Glu Pro Arg Arg Arg Arg Val Val Leu Gln His Val
 210 215 220

His Pro Pro His Phe Gly Pro Gly Gly His His Arg Arg His Arg Arg
 225 230 235 240

Leu Arg Gly Gln Arg Arg Ala Gln Gln Pro Ala His Leu Pro Asp Val
 245 250 255

Val Glu Gly Leu Glu Gln Val Val His His Ala Gln Leu Gln Arg Leu
 260 265 270

Leu Arg Arg His Gln Arg Arg Val Ala Arg His Gln Asp Asp Ala Arg
 275 280 285

Val Gly Val Glu Pro Pro Arg Leu Leu Arg Gln Leu Gln Pro Gly Gln
 290 295 300

Ala Arg Gln Val Leu Val Arg Asp Asp Glu Val His Trp Arg Leu Ala
 305 310 315 320

Gln Gln Ile Gln Arg Phe Leu Ala Ala Gly Arg Leu Asp Asp Leu Ala
 325 330 335

Leu Leu Ala His Gln Val Ala Glu Asp His Ala Asp Asp Gly Leu Val
 340 345 350

Val His His Gln Gln Pro Leu Ala Gln Leu Arg Arg Arg Val Arg Gly
 355 360 365

Gly Glu Glu Tyr Pro Glu His Gly Gly Arg Arg Gly His Val Ala Asn
 370 375 380

Gly Ala Arg Arg Arg Glu Ala Asn Arg Arg Val Val Leu Leu Gly Asp
 385 390 395 400

Ala Leu Ala His Arg Gln Pro Gln Ser Arg Ala Ala Leu Val Gly Gly
 405 410 415

Asp Val Gly Leu Glu Asp Ala Leu Ala Gln Leu His Gly Asp Ala Arg
 420 425 430

Ala Pro Val Arg His Val Gln His Arg Arg Gly Ala Leu Ala Arg Glu
 435 440 445

His His Leu His Ala Pro Ala Ala Gly His Gly Val Ala Arg Val Val
 450 455 460

Gln Gln Val Gln Arg Glu Leu Leu Asp Glu Pro Arg Val Ala Leu Asp
465 470 475 480

Gly Asp Val Leu Gly Arg Leu His Leu Gln Arg His Ala Arg Arg Val
485 490 495

Arg Val Asp Ala Glu Gln Leu Arg Arg Arg Pro Asp Gly Arg Arg Gln
500 505 510

Arg Gln Ala Leu Leu Arg Arg Val Pro Leu Ala Gly Glu Ala Asp Val
515 520 525

Val Gly Asp Asp Leu Phe His Pro Leu His Leu Arg Gln Gln Gln Pro
530 535 540

Gln Arg Ala Gly Gly Ala Thr Val Leu Arg Glu Gln Leu Arg Val Gly
545 550 555 560

Leu His Ala Gln Gln Gly Leu Pro Gln Leu Val Gln His Leu Arg Arg
565 570 575

Gln Leu Cys Arg Val Arg Ala Arg Thr Leu Gln Ser Arg Arg Arg Arg
580 585 590

Ala Arg Gly Arg Gln Ile Ser Arg Leu Gln Gly Leu Glu Gln Leu Gly
595 600 605

Glu Arg Arg Gly Cys Phe His Gly Pro Ser Met Ala Glu Ala Ala Val
610 615 620

Ser Thr Gln Arg Gly Arg Arg Asp Ser Val Asp Pro Gly Pro Trp Pro
625 630 635 640

Arg Gly Cys Ser Ser Arg Pro Leu Ala Pro Arg Pro Glu Arg Gly Arg
645 650 655

His His Arg Pro Arg Glu Gly Leu His Ala Ala Asp Gln Pro Val Pro
660 665 670

Leu Ala Gly Ala Pro Glu Asp Asp Arg Leu Arg Ala Gly Pro Leu His
675 680 685

Pro Ala Glu His Arg Gly Gly Ala Gly Arg His Gln Pro Leu Arg Arg
690 695 700

Pro His Arg Cys His Arg Gln Arg Gly Gly Gln Gly His Pro Pro Ser

705

710

715

720

Arg His Gly Arg Gly Ala Arg Ala Arg Arg Pro Leu Leu Ala Ala Asp
725 730 735

Cys Pro Val Gly Pro Arg Ala Gly Pro Ala
740 745

<210> 31

<211> 417

<212> PRT

<213> Angiococcus disciformis

<400> 31

Val Thr Ala Val Asp Asn Gly Leu Asp Ala Ile Ser Arg Thr Arg Glu
1 5 10 15

Leu Arg Pro Asp Val Val Leu Ala Asp Val Met Met Pro Gly Lys Ser
20 25 30

Gly Tyr Glu Val Cys Glu Ala Leu Lys Ser Asp Pro Ala Thr Gln Gly
35 40 45

Ile Pro Val Val Leu Leu Ala Gly Thr Phe Glu Ala Phe Asp Glu Asn
50 55 60

Arg Ala Arg Ala Ala Arg Ala Asp Asp His Ile Ala Lys Pro Phe Glu
65 70 75 80

Ser Gln Val Leu Leu Asp Lys Val Lys Ala Leu Val Gly Gln Lys Ser
85 90 95

Asn Thr Met Pro Ala Ser Ala Ala Thr Gln Val Arg Pro Pro Thr Ala
100 105 110

Gln Pro Ala Thr Pro Ala Pro Val Pro Pro Gly Pro Gln Pro Gly Ala
115 120 125

Pro Thr Ala Ala Arg Pro Gly Ile Pro Gln Gly Thr Gln Pro Gly Val
130 135 140

Pro Arg Pro Pro Gly Pro Gly Met Pro Pro Gly Ala Pro Gly Ala Pro
145 150 155 160

223

Arg Pro Pro Gly Pro Gly Met Pro Pro Gly Pro Gly Met Ala Arg Pro
 165 170 175

Pro Gly Pro Gly Met Pro Pro Pro Gly Ala Pro Gly Ala Pro Arg Pro
 180 185 190

Pro Gly Pro Gly Met Pro Pro Gly Pro Gly Met Ala Arg Pro Pro Gly
 195 200 205

Pro Gly Met Pro Pro Pro Gly Ala Pro Gly Ala Pro Arg Pro Pro Gly
 210 215 220

Pro Gly Met Pro Pro Pro Gly Ala Ala Ala Arg Pro Gly Arg Ala Ala
 225 230 235 240

Ser Ala Gly Cys Val Arg Ala Gly His Ala Ala Ala Ser Gly Arg Ala
 245 250 255

Arg Ala Ser Gly Trp Ala Gly Ala Gly Met Ala Arg Pro Gly Ile Pro
 260 265 270

Gln Gly Ser Gln Pro Gly Phe Pro Arg Pro Pro Gly Gly Ala Ala Pro
 275 280 285

Leu Pro Pro Ala Pro Gly Gly Ala Thr Gln Pro Ala Ala Arg Ala Arg
 290 295 300

Asp Pro Phe Gly Leu Gly Ala Pro Thr Ala Ala Pro Gln Ala Ala Gln
 305 310 315 320

Pro Ser Ile Ser Ile Glu Asp Ser Leu Pro Glu Pro Ser Gly Ala Glu
 325 330 335

Glu Ile Ser Leu Asp Met Gly Gly Pro Pro Thr Ala Ala Pro Thr Ala
 340 345 350

Val Ala Pro Thr Ala Ala Pro Ala Thr Ala Arg Pro Thr Ala Ala Asp
 355 360 365

Gly Gly Glu Ala Leu Leu Arg Glu Ala Leu Ser Lys Ala Ser Arg Glu
 370 375 380

Val Ile Glu Lys Ile Ala Trp Glu Val Val Pro Gln Leu Ala Glu Thr
 385 390 395 400

Ile Ile Arg Glu Glu Leu Glu Arg Leu Ile Lys Asp Arg Glu Thr Gln
 405 410 415

His

<210> 32

<211> 1102

<212> PRT

<213> Angiococcus disciformis

<400> 32

Met Glu Arg Asn Tyr Phe Arg Ala Glu Ala Thr Ser Asp Lys Pro Ala
1 5 10 15

Phe Ser Ile Val Leu Pro Pro Pro Asn Val Thr Gly Ser Leu His Ile
20 25 30

Gly His Ala Leu Thr Ala Thr Ile Gln Asp Ile Leu Ala Arg Trp Lys
35 40 45

Arg Met Ser Gly Phe Asn Thr Leu Trp Leu Pro Gly Thr Asp His Ala
50 55 60

Gly Ile Ala Thr Gln Met Val Val Glu Lys Glu Leu Lys Lys Thr Glu
65 70 75 80

Gly Lys Ser Arg His Asp Leu Gly Arg Ala Ala Phe Leu Glu Arg Val
85 90 95

Trp Glu Trp Lys Gly Lys Tyr Gly Ala Arg Ile Gly Glu Gln His Arg
100 105 110

Phe Leu Gly Ala Ser Leu Asp Trp Ser Arg Glu Arg Phe Thr Met Asp
115 120 125

Glu Gln Ser Ser Ala Ala Val Arg Glu Val Phe Val Arg Leu Tyr Glu
130 135 140

Glu Gly Leu Met Tyr Arg Ala Gln Lys Leu Ile Asn Trp Cys Pro Ser
145 150 155 160

Cys Arg Thr Ala Leu Ser Asp Leu Glu Val Glu His Glu Glu Ser Lys
165 170 175

Lys Gly Ser Ile Trp His Ile Arg Tyr Pro Val Lys Asp Ser Asp Arg
180 185 190

Thr Leu Thr Val Ala Thr Thr Arg Pro Glu Thr Met Leu Gly Asp Thr
 195 200 205

Ala Val Ala Ile His Pro Glu Asp Glu Arg Tyr Leu Gly Leu Ala Gly
 210 215 220

Lys Ser Val Lys Leu Pro Leu Thr Asp Arg Glu Ile Pro Ile Ile Ala
 225 230 235 240

Asp Ala Glu Leu Val Asp Pro Lys Phe Gly Thr Gly Val Val Lys Val
 245 250 255

Thr Pro Ala His Asp Phe Asn Asp Tyr Gln Thr Gly Leu Arg His Lys
 260 265 270

Leu Pro Met Leu Thr Ile Leu Asp Glu Ala Ala Arg Met Thr Lys Glu
 275 280 285

Thr Gly Lys Tyr Ala Gly Met Asp Arg Phe Glu Ala Arg Lys Gln Val
 290 295 300

Leu Ala Asp Leu Thr Glu Gln Gly Leu Leu Glu Lys Glu Glu Pro His
 305 310 315 320

Met Leu Asn Val Gly Thr Cys Gln Arg Cys Ala Thr Val Val Glu Pro
 325 330 335

Arg Leu Ser Pro Gln Trp Phe Ile Lys Ile Glu Pro Leu Ala Lys Pro
 340 345 350

Ala Ile Glu Ala Val Glu Gln Gly Arg Thr Lys Phe Val Pro Glu Ser
 355 360 365

Trp Thr Asn Thr Phe Phe His Trp Met Arg Asn Ile His Asp Trp Cys
 370 375 380

Val Ser Arg Gln Leu Trp Trp Gly His Gln Ile Pro Ala Tyr Tyr Cys
 385 390 395 400

Thr Ser Cys Ser Pro Arg Gln Gly Asp Asp Thr Asp Leu Pro Leu Asp
 405 410 415

Ala Pro Thr Val Lys Val Gly Gly Val Asp Phe Ala Arg Ala Glu Pro
 420 425 430

Ile Val Ala Arg Glu Gln Pro Lys Ser Cys Pro Lys Cys Gly Gly Ala

435

440

445

Ser Phe Ile Gln Asp Pro Asp Val Leu Asp Thr Trp Phe Ser Ser Gly
450 455 460

Leu Trp Pro Phe Ser Thr Leu Gly Trp Pro Arg Glu Thr Pro Glu Leu
465 470 475 480

Lys Thr Phe Tyr Pro Thr Ser Val Met Glu Thr Gly His Asp Ile Ile
485 490 495

Phe Phe Trp Val Ala Arg Met Met Met Met Gly Ile His Phe Met Gly
500 505 510

Asp Val Pro Phe Arg Thr Val Tyr Leu His Ala Met Val Arg Asp Glu
515 520 525

Lys Gly Gln Lys Met Ser Lys Thr Lys Gly Asn Val Ile Asp Pro Leu
530 535 540

Asp Ile Val Leu Gly Ala Pro Ala Asp Lys Leu Gln Pro Thr Leu Lys
545 550 555 560

Asn Arg Phe Pro Gln Gly Met Pro Pro Phe Gly Ala Asp Ala Leu Arg
565 570 575

Phe Thr Leu Ala Ser Leu Thr Gln Gln Gly Arg Asp Ile Lys Leu Ser
580 585 590

Met Asp Arg Leu Glu Gly Tyr Lys Ala Phe Gly Asn Lys Leu Trp Asn
595 600 605

Ala Ser Arg Phe Ala Leu Met Asn Met Gly Asp Phe Arg Met Glu Ala
610 615 620

Gly Pro Leu Asp Arg Ser Thr Leu Thr Leu Ala Asp Arg Trp Ile Leu
625 630 635 640

Ser Arg Leu Gln Arg Ala Thr Thr Glu Ala Arg Ala Leu Leu Glu Ser
645 650 655

Tyr Ser Phe Gly Glu Ala Ala Ser Thr Leu Tyr Gln Phe Leu Trp Ala
660 665 670

Glu Phe Cys Asp Trp Tyr Ile Glu Leu Ala Lys Gly Ser Leu Tyr Gly
675 680 685

227

Glu Asp Ala Ala Ala Lys Asp Ser Thr Arg Arg Val Leu Val Tyr Cys
 690 695 700

Leu Asp Arg Ile Leu Arg Leu Leu His Pro Phe Met Pro Phe Ile Thr
 705 710 715 720

Glu Glu Ile Trp Gln Lys Leu Pro Met Ala Arg Pro Thr Asp Ser Ile
 725 730 735

Met Ile Ala Pro Tyr Pro Glu Pro Asp Ala Ala His Val Asp Glu Ala
 740 745 750

Ala Glu Gly Glu Met Ala Pro Val Ile Ala Ala Ile Glu Gly Leu Arg
 755 760 765

Thr Ile Arg Gly Glu Ser Asn Leu Ser Pro Ala Thr Lys Val Lys Ala
 770 775 780

Val Val Gln Ser Ala Asp Val Arg Thr Arg Glu Leu Leu Glu Arg Trp
 785 790 795 800

Arg Gly Tyr Leu Met Pro Leu Ala Gly Leu Ser Glu Val Thr Val Gly
 805 810 815

Ala Pro Gly Ala Lys Pro Pro Gln Ala Ala Ala Phe Val Gly Ser Asn
 820 825 830

Leu Glu Ile Tyr Val Pro Leu Ala Gly Leu Val Asp Leu Asp Ala Glu
 835 840 845

Arg Glu Arg Leu Lys Lys Glu Ile Ala Arg Ala Glu Gln Glu Met Ala
 850 855 860

Gly Val Leu Arg Lys Leu Glu Asn Pro Asn Phe Val Ala Lys Ala Pro
 865 870 875 880

Pro Asp Val Val Glu Lys Asp Arg Ala Arg Val Glu Glu Leu Lys Glu
 885 890 895

Arg Lys Ala Lys Leu Gln Asp His Leu Gln Arg Ile Ala Pro Glu Pro
 900 905 910

Ala Met Pro Glu Thr Pro Pro Ser Glu Ser Thr Thr Pro Glu Ser Ser
 915 920 925

Thr Pro Thr Ala Glu Glu Gly Ala Ser Pro Ser Ala Pro Ala Gln Val
 930 935 940

Lys Val Ala Pro Asn Gln Glu Glu Lys Gly Gly Val Asp Leu Gly Gln
945 950 955 960

Glu Leu Lys Gly Glu Leu Gly Glu Thr Gly Thr Ala Ser Glu Ala Gly
965 970 975

Asp Thr Gln Val Gln Asp Ala Leu Ala Lys Leu Arg Glu Gly Thr Lys
980 985 990

Glu Gly Leu Ser Pro Ser Asp His His Asp Leu Gly Val Ala Tyr Met
995 1000 1005

Thr Met Gly Leu Val Asp Asp Ala Met Arg Glu Phe Asn Arg Ala
1010 1015 1020

Lys Glu Gly Gly Asp Thr Arg Glu Thr Pro Ala Glu Gly Ala Glu
1025 1030 1035

Ala Glu Glu Gly Glu Ser Gly Pro Val Lys Ala His Gly Asp Met
1040 1045 1050

Leu Pro Thr Arg Glu Val Thr Gln Gly Glu Pro Val Arg Ser Ala
1055 1060 1065

Ser Lys Lys Lys Ala Thr Ala Pro Gly Lys Val His Gly Asp Lys
1070 1075 1080

Leu Ala Pro Ala Pro Thr Thr Glu Glu Ala Glu Ala Pro Ala Val
1085 1090 1095

Lys Lys Ala Ala
1100

<210> 33

<211> 21953

<212> DNA

<213> *Angiococcus disciformis*

<220>

<221> CDS

<222> (1) .. (7878)

<223> tubC

<220>

<221> CDS

<222> (18408)..(21953)

<223> tubE

<400> 33

ttg agc gcc gga gcg ttg ctc gcc cac gcc gcg agc ctg ggt gta cgc	48
Leu Ser Ala Gly Ala Leu Leu Ala His Ala Ala Ser Leu Gly Val Arg	
1 5 10 15	
ctg tgg gta gag ggc gag cga ctg cgc ttc cag gct ccc cca ggg gtg	96
Leu Trp Val Glu Gly Glu Arg Leu Arg Phe Gln Ala Pro Pro Gly Val	
20 25 30	
atg acg ccc gag ctg cag tcc cgc ctg gga ggc gct cgc cat gag ctc	144
Met Thr Pro Glu Leu Gln Ser Arg Leu Gly Gly Ala Arg His Glu Leu	
35 40 45	
atc gca ctg cta cgg cag ctc cag ccc tcc tct cag gga ggc agc ctc	192
Ile Ala Leu Leu Arg Gln Leu Gln Pro Ser Ser Gln Gly Gly Ser Leu	
50 55 60	
ctg gcc cca gtg gct cga aac ggg cgg ctg gcg ctc tcc ttc gcc cag	240
Leu Ala Pro Val Ala Arg Asn Gly Arg Leu Ala Leu Ser Phe Ala Gln	
65 70 75 80	
cag cgg ctg tgg ttc cag gag cag ctc cac cct gag gcc ccg gcc aac	288
Gln Arg Leu Trp Phe Gln Glu Gln Leu His Pro Glu Ala Pro Ala Asn	
85 90 95	
aac ctc acg ggg gcc gtg gtg ttc acc ggc ccg ctg cac gtg gcc gcg	336
Asn Leu Thr Gly Ala Val Val Phe Thr Gly Pro Leu His Val Ala Ala	
100 105 110	
ctg ctc ggg gcg gtg gct gcg ctg gtg agg cga cac gaa gcg ctg cgc	384
Leu Leu Gly Ala Val Ala Ala Leu Val Arg Arg His Glu Ala Leu Arg	
115 120 125	
acc acc ctg ggc gag gag ggc gga gtc ccc tat tcg ctc atc ggc gag	432
Thr Thr Leu Gly Glu Glu Gly Gly Val Pro Tyr Ser Leu Ile Gly Glu	
130 135 140	
ccg tgg cag ccg gcg ctg gag gtg gaa gcg cta ccg ggt gcc acc gtg	480
Pro Trp Gln Pro Ala Leu Glu Val Glu Ala Leu Pro Gly Ala Thr Val	
145 150 155 160	
ggc gag cgg ctg gag cag gcc cga gag gtg gcc ctg gcc gag tcc cgg	528
Gly Glu Arg Leu Glu Gln Ala Arg Glu Val Ala Leu Ala Glu Ser Arg	
165 170 175	
cgc cgc ttc gcc ctg gag acg gag ccc cac ctg cgc gtg cgc ctg ctg	576
Arg Arg Phe Ala Leu Glu Thr Glu Pro His Leu Arg Val Arg Leu Leu	
180 185 190	
cgc ctg gcc gag cag cag cac gtg ctg gtg ctc agc ctc cac cac atc	624
Arg Leu Ala Glu Gln Gln His Val Leu Val Leu Ser Leu His His Ile	

230

195	200	205	
gcg gcg gat gga gtc ggg ctg cag gtg ctc gaa cag gag ctg gcg gcg Ala Ala Asp Gly Val Gly Leu Gln Val Leu Glu Gln Glu Leu Ala Ala 210 215 220			672
ctc tac ggc gcg ctg agc gct ggg gcc gag ccg cgg cta cct ccc ctg Leu Tyr Gly Ala Leu Ser Ala Gly Ala Glu Pro Arg Leu Pro Pro Leu 225 230 235 240			720
ccc ctg cag gtg gcg gac ctt gcc gac tgg cag cgc cgg tgg gtg gag Pro Leu Gln Val Ala Asp Leu Ala Asp Trp Gln Arg Arg Trp Val Glu 245 250 255			768
gga gag gag tac cag gtc cag ctc gct tac tgg cgg cga cag ctc gcc Gly Glu Glu Tyr Gln Val Gln Leu Ala Tyr Trp Arg Arg Gln Leu Ala 260 265 270			816
ggg ctc acg ccc ctg gag gta cca gga gat cat cct cgc ccg cgt atc Gly Leu Thr Pro Leu Glu Val Pro Gly Asp His Pro Arg Pro Arg Ile 275 280 285			864
ccc tcc atg aga ggc gcg gag gta cgg gcg ccc ctg ctc tcc gct ccg Pro Ser Met Arg Gly Ala Glu Val Arg Ala Pro Leu Leu Ser Ala Pro 290 295 300			912
caa gcc cag gtg ctg cgc gcg ctg ggg caa ggc gag ggc gcc acc ctc Gln Ala Gln Val Leu Arg Ala Leu Gly Gln Gly Glu Gly Ala Thr Leu 305 310 315 320			960
tac atg acg ctc ctg gcc gcc ctg ggc gtg ctg ctg caa cgg tgg acg Tyr Met Thr Leu Leu Ala Ala Leu Gly Val Leu Leu Gln Arg Trp Thr 325 330 335			1008
ggc cag cac gat atg gcg gtg ggc agc gcc gcc gcc aac cgc aac cgc Gly Gln His Asp Met Ala Val Gly Ser Ala Ala Ala Asn Arg Asn Arg 340 345 350			1056
ccg ggc ttg gag ggc atc ctc ggc ttc ctc ctc aac atc gtc ctg ctg Pro Gly Leu Glu Gly Ile Leu Gly Phe Leu Leu Asn Ile Val Leu Leu 355 360 365			1104
cgg ctc gac ttg aga ggc cgc ccc cgc ttc cgc gag ttg ctg cgt cag Arg Leu Asp Leu Arg Gly Arg Pro Arg Phe Arg Glu Leu Leu Arg Gln 370 375 380			1152
gca cgc agg gtg tgt gtg gag gcc tat gcc cat cag gag ttg ccc ttc Ala Arg Arg Val Cys Val Glu Ala Tyr Ala His Gln Glu Leu Pro Phe 385 390 395 400			1200
gag cac ctg gtg gag gcg ctc cag cca ggc agc gag cgc ggg gac agc Glu His Leu Val Glu Ala Leu Gln Pro Gly Ser Glu Arg Gly Asp Ser 405 410 415			1248
tgg ctc tat cgc gtg gcg ctc gcg gtg agc gac acg cca tgg atg ccc Ser Leu Tyr Arg Val Ala Leu Ala Val Ser Asp Thr Pro Trp Met Pro 420 425 430			1296
ccc cac ggg ctg aag ctg gag ggg gtg caa gcg cag ccg ctc gac ttt Gly His Gly Leu Lys Leu Glu Gly Val Gln Ala Gln Pro Leu Asp Phe 435 440 445			1344
ccc cgg ggg gtg ctg gac ctg gat ctc cac ctg tgg gtg tat gac acc			1392

Pro Arg Gly Val Leu Asp Leu Asp Leu His Leu Trp Val Tyr Asp Thr	
450 455 460	
ggg gag ggg ctc acc ggg cgg ttg gag tac gcc gtg gat ctg tat gag	1440
Gly Glu Gly Leu Thr Gly Arg Leu Glu Tyr Ala Val Asp Leu Tyr Glu	
465 470 475 480	
gag ccc acg gca agg cgg ctc ctg gag ggc ttc cgc cag gtg ctg gag	1488
Glu Pro Thr Ala Arg Arg Leu Leu Glu Gly Phe Arg Gln Val Leu Glu	
485 490 495	
gcg gtg gtg gag gcg ccg gat cgg ccc gtg ccg gag ctg ccg gtg ctg	1536
Ala Val Val Glu Ala Pro Asp Arg Pro Val Pro Glu Leu Pro Val Leu	
500 505 510	
gga gag cag gag cgc cac cag gtg ctc tcc ggg tgg aac cgc aca cag	1584
Gly Glu Gln Glu Arg His Gln Val Leu Ser Gly Trp Asn Arg Thr Gln	
515 520 525	
cgc ccc tat ccg cga gag gcc agc gtt cac ggc ctc ttt cag cag cgt	1632
Arg Pro Tyr Pro Arg Glu Ala Ser Val His Gly Leu Phe Gln Gln Arg	
530 535 540	
gcc ctc cag gcc ccg cgg gcc gtg gca gtg gtg tat ggc gag cgc agc	1680
Ala Leu Gln Ala Pro Arg Ala Val Ala Val Val Tyr Gly Glu Arg Ser	
545 550 555 560	
ctc acc tac ggc gag ttg gcg gag cgg gcc cgc ggg ctc gca cag ggt	1728
Leu Thr Tyr Gly Glu Leu Ala Glu Arg Ala Arg Gly Leu Ala Gln Gly	
565 570 575	
ctg gtg gcc cga ggc gtg cgg cga gga gac ctg gtc gcc ctt cgg ctg	1776
Leu Val Ala Arg Gly Val Arg Arg Gly Asp Leu Val Ala Leu Arg Leu	
580 585 590	
gag cgc tcg ccg gag cag gtg gag tcc atg ctg gcc gta ctc caa gcc	1824
Glu Arg Ser Pro Glu Gln Val Glu Ser Met Leu Ala Val Leu Gln Ala	
595 600 605	
ggg gcc gcc tac gtg ccg ctg gat cct tcc tac ccc gta cag cgc cag	1872
Gly Ala Ala Tyr Val Pro Leu Asp Pro Ser Tyr Pro Val Gln Arg Gln	
610 615 620	
gag ttc atg ctc cag gac agc ggc gcg cgg ttg ctc gtg cac agc ggc	1920
Glu Phe Met Leu Gln Asp Ser Gly Ala Arg Leu Leu Val His Ser Gly	
625 630 635 640	
ccg ctg ccc ttc gct ccc cag ggc tgc gcc acg ctg gat ctc cag gca	1968
Pro Leu Pro Phe Ala Pro Gln Gly Cys Ala Thr Leu Asp Leu Gln Ala	
645 650 655	
tgg cac ccc gct ccc tcc gac ggt ggc gag ccc ctg ccc cag tgc agc	2016
Trp His Pro Ala Pro Ser Asp Gly Gly Glu Pro Leu Pro Gln Cys Ser	
660 665 670	
gga gaa gat ctg gcg tat gtc atc tat acg tcc ggc tcc acc ggc cag	2064
Gly Glu Asp Leu Ala Tyr Val Ile Tyr Thr Ser Gly Ser Thr Gly Gln	
675 680 685	
ccc aag ggc gtg gcc gtg tgc cac cgg gcc atg acc cgc ctg gtg tgc	2112
Pro Lys Gly Val Ala Val Cys His Arg Ala Met Thr Arg Leu Val Cys	
690 695 700	

232

aac acg gac tac gtt cag ctc ggg cca gag gat cgg gtc gcc cag gct	2160
Asn Thr Asp Tyr Val Gln Leu Gly Pro Glu Asp Arg Val Ala Gln Ala	
705 710 715 720	
tcc aat gcc tcc ttt gac gcg gcc acc ttc gag gtc tgg ggc gcg ctg	2208
Ser Asn Ala Ser Phe Asp Ala Ala Thr Phe Glu Val Trp Gly Ala Leu	
725 730 735	
ctc aat ggg gca cgg ctg gtg ggg ctc gcc acg gag gag gcc att cag	2256
Leu Asn Gly Ala Arg Leu Val Gly Leu Ala Thr Glu Glu Ala Ile Gln	
740 745 750	
gcg cgc cgg ctg gcc gag gtc ctg cgc gag cag cgg atc tcc gtg ctc	2304
Ala Arg Arg Leu Ala Glu Val Leu Arg Glu Gln Arg Ile Ser Val Leu	
755 760 765	
ttc gtc acc acc gcc ctc ttc aac cac gtg gcc cgg gag caa ccc cag	2352
Phe Val Thr Thr Ala Leu Phe Asn His Val Ala Arg Glu Gln Pro Gln	
770 775 780	
gcc ttc agc acc ctg cgc tac ctg ctc ttt ggc ggc gag gca gtg gat	2400
Ala Phe Ser Thr Leu Arg Tyr Leu Leu Phe Gly Gly Glu Ala Val Asp	
785 790 795 800	
gcc tcc agt gtg agg cgg gtg ctg aaa cag ggc gct cca gga cac ctt	2448
Ala Ser Ser Val Arg Arg Val Leu Lys Gln Gly Ala Pro Gly His Leu	
805 810 815	
cta cac gtg tac ggc ccc acg gag aac acc acc ttc tcc acc gcc tgg	2496
Leu His Val Tyr Gly Pro Thr Glu Asn Thr Thr Phe Ser Thr Ala Trp	
820 825 830	
cgg gtg gaa cac ctg gcc gag cag gcg cat acc gtg ccc atg ggt cat	2544
Arg Val Glu His Leu Ala Glu Gln Ala His Thr Val Pro Met Gly His	
835 840 845	
ccc atc gcc aac tcg cgg ctg cac gtg ctg gac gaa gcg ctg cag ccg	2592
Pro Ile Ala Asn Ser Arg Leu His Val Leu Asp Glu Ala Leu Gln Pro	
850 855 860	
gtg ccg gtg ggc gcc atg ggc gag gta tac ctg gga ggt gat ggg ctg	2640
Val Pro Val Gly Ala Met Gly Glu Val Tyr Leu Gly Gly Asp Gly Leu	
865 870 875 880	
gcg ctg ggc tac tgg cgc cac ccc gag gcc acc gcg gag cgc ttc gtg	2688
Ala Leu Gly Tyr Trp Arg His Pro Glu Ala Thr Ala Glu Arg Phe Val	
885 890 895	
ccg gat ccg cac ggc ctg gag ccc gga ggc cgg ctc tac cgc acg gga	2736
Pro Asp Pro His Gly Leu Glu Pro Gly Gly Arg Leu Tyr Arg Thr Gly	
900 905 910	
gat ctg gcg cgg cgg cag gcg gat ggg gcc gtg gtg ttc gct ggc cgc	2784
Asp Leu Ala Arg Arg Gln Ala Asp Gly Ala Val Val Phe Ala Gly Arg	
915 920 925	
gtg gac agg cag gtc aag ctg cga ggc ttc cgc gtg gag ccg gcg gag	2832
Val Asp Arg Gln Val Lys Leu Arg Gly Phe Arg Val Glu Pro Ala Glu	
930 935 940	
atc gag tcg cac ctg tgc gag cac tcg gag gtg tcc gcg gcg gtg gtg	2880
Ile Glu Ser His Leu Cys Glu His Ser Glu Val Ser Ala Ala Val Val	
945 950 955 960	

gag ctg cgg ggc gaa ggg gcc ctg cgg cgg ctg gtc gcc tat gtc gtc	2928
Glu Leu Arg Gly Glu Gly Ala Leu Arg Arg Leu Val Ala Tyr Val Val	
965 970 975	
ccc cgc gcc ggt ggc cgc cca ggt gcc gag gag ctg cgc acc ttc ctc	2976
Pro Arg Ala Gly Gly Arg Pro Gly Ala Glu Glu Leu Arg Thr Phe Leu	
980 985 990	
cgc acc cga ctg ccc gag tac atg ctg ccc gcc agc ttc tcc ctg ctg	3024
Arg Thr Arg Leu Pro Glu Tyr Met Leu Pro Ala Ser Phe Ser Leu Leu	
995 1000 1005	
gag gca ctg ccc ctc acc ccc aac ggc aag gtg gat cgc tcg gcg	3069
Glu Ala Leu Pro Leu Thr Pro Asn Gly Lys Val Asp Arg Ser Ala	
1010 1015 1020	
cta ccc gag agc ttc gag gag gcc tcc cgc gag cag gct ccg gtg	3114
Leu Pro Glu Ser Phe Glu Glu Ala Ser Arg Glu Gln Ala Pro Val	
1025 1030 1035	
gta cct cca cga ggc ccc gtg gag gcg ctg ctg gtg gac atc tgg	3159
Val Pro Pro Arg Gly Pro Val Glu Ala Leu Leu Val Asp Ile Trp	
1040 1045 1050	
agg gag gtg ctc ggc acc caa cgg gta tcc gtg cac gac gac ttc	3204
Arg Glu Val Leu Gly Thr Gln Arg Val Ser Val His Asp Asp Phe	
1055 1060 1065	
ttc gac ctg ggc gga cac tcc ctg ctg gcc acg cgg gtt gtc tcc	3249
Phe Asp Leu Gly Gly His Ser Leu Leu Ala Thr Arg Val Val Ser	
1070 1075 1080	
cgc ttg cga gag gcg ctg cag gtg gag ttg ccg ctg cgc acc ctc	3294
Arg Leu Arg Glu Ala Leu Gln Val Glu Leu Pro Leu Arg Thr Leu	
1085 1090 1095	
ttc gag gcc cct cag ctc tcg gcg ctg gcg gcc cag gtg gag gtc	3339
Phe Glu Ala Pro Gln Leu Ser Ala Leu Ala Ala Gln Val Glu Val	
1100 1105 1110	
ctg ctc ggc cac cgg cag ctc cgc ccc cct ccg ctc gtg ccc gct	3384
Leu Leu Gly His Arg Gln Leu Arg Pro Pro Pro Leu Val Pro Ala	
1115 1120 1125	
gta cgt cca cca gag ctg ccc ctc tcc ttc gcc cag cag cgg ctg	3429
Val Arg Pro Pro Glu Leu Pro Leu Ser Phe Ala Gln Gln Arg Leu	
1130 1135 1140	
tgg ttc ctc cag cag ctc gcg cca caa agc acc gcc tac cag atc	3474
Trp Phe Leu Gln Gln Leu Ala Pro Gln Ser Thr Ala Tyr Gln Ile	
1145 1150 1155	
ctg gat gcg tgg cat gtg cgg ggc cgg gtg gac gtg ggc gcc ctg	3519
Leu Asp Ala Trp His Val Arg Gly Arg Val Asp Val Gly Ala Leu	
1160 1165 1170	
gag cgc gcc ctg gag cag ctc gtg cgc cgg cat gag gcg ctg cgc	3564
Glu Arg Ala Leu Glu Gln Leu Val Arg Arg His Glu Ala Leu Arg	
1175 1180 1185	
acc act ttc gag cct ggc ggc gac ggc gtg cct cgg cag cgc atc	3609
Thr Thr Phe Glu Pro Gly Gly Asp Gly Val Pro Arg Gln Arg Ile	

1190	1195	1200	
cac gcg ccg gct ccg gtc ccg ctg cgg cag gtg gac ctg cgc tcc			3654
His Ala Pro Ala Pro Val Pro Leu Arg Gln Val Asp Leu Arg Ser			
1205	1210	1215	
cat ggc atc gcc gcg cgg gag gag gcc ttg cgg tgg atg cgc gag			3699
His Gly Ile Ala Ala Arg Glu Glu Ala Leu Arg Trp Met Arg Glu			
1220	1225	1230	
cag gcg ctc cgg cca ctg gag ctg gac aag ggg ccg ctg ctg cgc			3744
Gln Ala Leu Arg Pro Leu Glu Leu Asp Lys Gly Pro Leu Leu Arg			
1235	1240	1245	
gtg tcc ctg ctg cgc ttg gaa gag gca cag tcc ctg ctc ttc ctc			3789
Val Ser Leu Leu Arg Leu Glu Glu Ala Gln Ser Leu Leu Phe Leu			
1250	1255	1260	
gag ctg cac cat atc gtg ggc gat ggc tgg tcg ctg agc gtt tgg			3834
Glu Leu His His Ile Val Gly Asp Gly Trp Ser Leu Ser Val Trp			
1265	1270	1275	
agc cga gag ctg tcg cac ctg tac gag gcg gcg ctc cat ggc gct			3879
Ser Arg Glu Leu Ser His Leu Tyr Glu Ala Ala Leu His Gly Ala			
1280	1285	1290	
gag ccg ggg ttg gcg ccg ctg ccc gtg cag tac gcg gac ttc gcg			3924
Glu Pro Gly Leu Ala Pro Leu Pro Val Gln Tyr Ala Asp Phe Ala			
1295	1300	1305	
ctg tgg cag cgc ggc tgg ttg caa ggg ccg gtg ctg cgg gag gag			3969
Leu Trp Gln Arg Gly Trp Leu Gln Gly Pro Val Leu Arg Glu Glu			
1310	1315	1320	
ctg acc tgg tgg cgg gag ccg ctc gcg cgc ctg gcc ccg ctg cgc			4014
Leu Thr Trp Trp Arg Glu Arg Leu Ala Arg Leu Ala Pro Leu Arg			
1325	1330	1335	
ctg ccc gcg gac cat gcc cgc ccg gaa gtg cag cgc ttc aac ggc			4059
Leu Pro Ala Asp His Ala Arg Pro Glu Val Gln Arg Phe Asn Gly			
1340	1345	1350	
gcg acg tac cgc ttc acg ctc ccg ggc gta cgg gtg cag gcg ctg			4104
Ala Thr Tyr Arg Phe Thr Leu Pro Gly Val Arg Val Gln Ala Leu			
1355	1360	1365	
cgg cgg ctc ggg cat gag cac ggg gcc aca ctc ttc atg gtg ttg			4149
Arg Arg Leu Gly His Glu His Gly Ala Thr Leu Phe Met Val Leu			
1370	1375	1380	
ctc gcc ggc ttc aat gcg ctg ctg gcg cgc tac acc ggg cag acg			4194
Leu Ala Gly Phe Asn Ala Leu Leu Ala Arg Tyr Thr Gly Gln Thr			
1385	1390	1395	
gac atc gcc att ggc gcg ccc atc gcc aac cgc aca cgt ggg gag			4239
Asp Ile Ala Ile Gly Ala Pro Ile Ala Asn Arg Thr Arg Gly Glu			
1400	1405	1410	
gtg gag ggc ctc atc ggc ttc ttc gtc aac acc ctg gtg ctg cgc			4284
Val Glu Gly Leu Ile Gly Phe Phe Val Asn Thr Leu Val Leu Arg			
1415	1420	1425	
acc cgc ctg gaa ggc aac ccc tcc ttc ctc gag ctg ttg agg cgg			4329

Thr Arg 1430	Leu Glu Gly Asn 1435	Pro Ser Phe Leu Glu Leu Arg Arg 1440	
gtg agg 1445	gag aca acg cta 1450	gag gcc tac gcc cac 1455	gag ctg cct 4374
Val Arg 1445	Glu Thr Thr Leu 1450	Ala Tyr Ala His 1455	Gln Glu Leu Pro
ttc gag 1460	agg ctc gtc gag 1465	gag ttg cag ccc gag 1470	cag gcc aac 4419
Phe Glu 1460	Arg Leu Val Glu 1465	Leu Gln Pro Glu Arg 1470	Gln Ala Asn
cag aac 1475	ccg ctg gtg cag 1480	gtg ctg ctc gcc ctg 1485	cag aac gcg ccc 4464
Gln Asn 1475	Pro Leu Val Gln 1480	Val Leu Leu Ala Leu 1485	Gln Asn Ala Pro
cgc gag 1490	ccg ttg cgg ctg 1495	gag ggc ctg gag gcc 1500	cac ctg gag 4509
Arg Glu 1490	Pro Leu Arg Leu 1495	Ala Gly Leu Glu Ala 1500	Glu His Leu Glu
tac ctg 1505	gtg gcc acc acg 1510	cgc ttc gac ctc gag 1515	cac ctg tgg 4554
Tyr Leu 1505	Val Ala Thr Thr 1510	Phe Asp Leu Glu 1515	Leu His Leu Trp
gag gag 1520	gaa gag ggg ctg 1525	agc tgc atc gcc gtg 1530	tat gac cgc gcg 4599
Glu Glu 1520	Glu Glu Gly Leu 1525	Ser Cys Ile Ala Val 1530	Tyr Asp Arg Ala
ctg tac 1535	ggc gcg ggg acg 1540	gtg gag cgg ctg gtg 1545	ggc gcc tgg tgc 4644
Leu Tyr 1535	Gly Ala Gly Thr 1540	Val Glu Arg Leu Val 1545	Gly Ala Trp Cys
acg ctg 1550	ctg gag ggc gtg 1555	gcc gag ctc cgg gcg 1560	cgc cgc gtg gcg 4689
Thr Leu 1550	Leu Glu Gly Val 1555	Ala Glu Leu Pro Ala 1560	Arg Arg Val Ala
gag ctg 1565	ccc ctg gtg cct 1570	gcc cgg gag ctg cga 1575	cag gtt ccg ccg 4734
Glu Leu 1565	Pro Leu Val Pro 1570	Ala Arg Glu Leu Arg 1575	Gln Val Pro Pro
ccc tcc 1580	ccc gcc gcg gcg 1585	atc gaa aca agc atc 1590	ggc gcc cgc ttc 4779
Pro Ser 1580	Pro Ala Ala Ala 1585	Ile Glu Thr Ser Ile 1590	Gly Ala Arg Phe
tcg gag 1595	gtg gcg cgg cgg 1600	cag cca ggc gcc acg 1605	gcc gtc acc caa 4824
Ser Glu 1595	Val Ala Arg Arg 1600	Gln Pro Gly Ala Thr 1605	Ala Val Thr Gln
ggc ggg 1610	cgg cac ctc acc 1615	tat gca gag ctg gag 1620	gag cgc tcg gaa 4869
Gly Gly 1610	Arg His Leu Thr 1615	Tyr Ala Glu Leu Glu 1620	Glu Arg Ser Glu
cgc ctg 1625	gcc cgc tac ctg 1630	gcc tgg ctg ggc gtg 1635	cgc gcg ggc gac 4914
Arg Leu 1625	Ala Arg Tyr Leu 1630	Ala Trp Leu Gly Val 1635	Arg Ala Gly Asp
aga gtg 1640	ggg ttg gcc acc 1645	gag cgc acg ctg gag 1650	cgg atc atc agc 4959
Arg Val 1640	Gly Leu Ala Thr 1645	Glu Arg Thr Leu Glu 1650	Arg Ile Ile Ser
ctg cta 1655	ggc atc ctc aag 1660	gcg gga gcc gcc tac 1665	gta ccg ctg gat 5004
Leu Leu 1655	Gly Ile Leu Lys 1660	Ala Gly Ala Ala Tyr 1665	Val Pro Leu Asp

236

gtg cgc cag cca gcg cgt cgg	ctg agc ctg ctg gtg cag gcc gcg	5049
Val Arg Gln Pro Ala Arg Arg	Leu Ser Leu Leu Val Gln Ala Ala	
1670	1675 1680	
ggc gtg cgc acc gtc atc gcc	gag gag cag gct cgc acg gtg ctg	5094
Gly Val Arg Thr Val Ile Ala	Glu Glu Gln Ala Arg Thr Val Leu	
1685	1690 1695	
tca ggc ctg ggg cag cca ctg	acc ctg gtc gat gcc gcc cag gag	5139
Ser Gly Leu Gly Gln Pro Leu	Thr Leu Val Asp Ala Ala Gln Glu	
1700	1705 1710	
ccg gcc tcc gcg cag cag gtc	ccc gcg ctc ggg ccg gag cgc tcg	5184
Pro Ala Ser Ala Gln Gln Val	Pro Ala Leu Gly Pro Glu Arg Ser	
1715	1720 1725	
ctg ggc ggg gac atg ctc gcc	tac gtg ctt ttc acc tcg ggc tcc	5229
Leu Gly Gly Asp Met Leu Ala	Tyr Val Leu Phe Thr Ser Gly Ser	
1730	1735 1740	
acc ggc gag ccc aag ggc gtc	tgc att ccc cac cgt gcc gtg ctg	5274
Thr Gly Glu Pro Lys Gly Val	Cys Ile Pro His Arg Ala Val Leu	
1745	1750 1755	
cgg ctg atc cat gag ccc tcc	tac gtc cag ctc tcg ccg cgg gag	5319
Arg Leu Ile His Glu Pro Ser	Tyr Val Gln Leu Ser Pro Arg Glu	
1760	1765 1770	
gtc atg ctg cac tac gca ccg	ctg gag ttc gat gcc tcc acc ttc	5364
Val Met Leu His Tyr Ala Pro	Leu Glu Phe Asp Ala Ser Thr Phe	
1775	1780 1785	
gag gtc tgg ggg gcg ctg ctc	aac ggg gca agg ctc gtg ctg gtc	5409
Glu Val Trp Gly Ala Leu Leu	Asn Gly Ala Arg Leu Val Leu Val	
1790	1795 1800	
cct ccg gaa cag cag tcg ctc	gag agc ctg gga cag gag ctg agc	5454
Pro Pro Glu Gln Gln Ser Leu	Glu Ser Leu Gly Gln Glu Leu Ser	
1805	1810 1815	
acc cag ggc gtc acc gtg ctg	tgg ctc acc gcc ggg ctc ttc cgg	5499
Thr Gln Gly Val Thr Val Leu	Trp Leu Thr Ala Gly Leu Phe Arg	
1820	1825 1830	
ctc atg gtg gag gag cag ctc	aag agc ctg cgc ggc gtg cgc cag	5544
Leu Met Val Glu Glu Gln Leu	Lys Ser Leu Arg Gly Val Arg Gln	
1835	1840 1845	
ctg ctc gcc gga ggg gac gtg	ctg ccc atg ccg cag gta cgc cgg	5589
Leu Leu Ala Gly Gly Asp Val	Leu Pro Met Pro Gln Val Arg Arg	
1850	1855 1860	
ctg cgc gag gcc ctc ccc gag	tgc cag ctc atc aac ggc tac ggc	5634
Leu Arg Glu Ala Leu Pro Glu	Cys Gln Leu Ile Asn Gly Tyr Gly	
1865	1870 1875	
ccc acc gag agc tgc acc ttc	acc tgc tgt cac cgc gtg ggc agc	5679
Pro Thr Glu Ser Cys Thr Phe	Thr Cys Cys His Arg Val Gly Ser	
1880	1885 1890	
ccc cag gag ctg ggc ggc tcc	gtg ccc atc ggc acc ccc atc gac	5724
Pro Glu Glu Leu Gly Gly Ser	Val Pro Ile Gly Thr Pro Ile Asp	
1895	1900 1905	

ctg ggc tgg gtg agc gtg gtg gac gag cgg ctc cag ccc gtg ccg Leu Gly Trp Val Ser Val Val Asp Glu Arg Leu Gln Pro Val Pro 1910 1915 1920	5769
gac gga gcc ccg ggc gag ctg ctc gtg ggc ggc ccc ggg ctg gct Asp Gly Ala Pro Gly Glu Leu Leu Val Gly Gly Pro Gly Leu Ala 1925 1930 1935	5814
tgg ggg tac ctg caa cac ccg gag ctc acc gcc gag cgc ttc atc Trp Gly Tyr Leu Gln His Pro Glu Leu Thr Ala Glu Arg Phe Ile 1940 1945 1950	5859
ccc gat ccg ctc agc cgg acg ccg ggg gct cgc gtc tat cgg acg Pro Asp Pro Leu Ser Arg Thr Pro Gly Ala Arg Val Tyr Arg Thr 1955 1960 1965	5904
gga gac ctg gtg cgg cgc agg gag gat ggg acg ctg gag ttc ctg Gly Asp Leu Val Arg Arg Arg Glu Asp Gly Thr Leu Glu Phe Leu 1970 1975 1980	5949
ggc cgc gtg gac cac cag ctc aag gta cgc ggc ttc cgc atc gag Gly Arg Val Asp His Gln Leu Lys Val Arg Gly Phe Arg Ile Glu 1985 1990 1995	5994
ccg ggc gag gtg gag gcc gcg gtg ctc acc cac ccg gcg gtg cag Pro Gly Glu Val Glu Ala Ala Val Leu Thr His Pro Ala Val Gln 2000 2005 2010	6039
tcc gcg gtg gtg gtg ggg cgc gag ggc ccc ggc ggc aag gag ctg Ser Ala Val Val Val Gly Arg Glu Gly Pro Gly Gly Lys Glu Leu 2015 2020 2025	6084
gtc tgc tac gcc gtg ccg cgg gtg gag agc tcg gag cag ggc tcg Val Cys Tyr Ala Val Pro Arg Val Glu Ser Ser Glu Gln Gly Ser 2030 2035 2040	6129
cag cag gag cag cgg ctg gta cac gag tgg gag tcc gtg ttc gac Gln Gln Glu Gln Arg Leu Val His Glu Trp Glu Ser Val Phe Asp 2045 2050 2055	6174
ggg cac atg tac cgc gaa gcc ccc gtg ggc gga gag ccg acc ttc Gly His Met Tyr Arg Glu Ala Pro Val Gly Gly Glu Pro Thr Phe 2060 2065 2070	6219
aac atc gtc ggc tgg aag agc agc tac acc ggc caa ccc gtg gcc Asn Ile Val Gly Trp Lys Ser Ser Tyr Thr Gly Gln Pro Val Ala 2075 2080 2085	6264
gtg gag gag atg agg gac tgg ctg cgc cac cgg gtg gag cgc gtg Val Glu Glu Met Arg Asp Trp Leu Arg His Arg Val Glu Arg Val 2090 2095 2100	6309
cgc ggg ctc agg cca cga cgg atc ctc gaa gtg ggc tgt ggt acg Arg Gly Leu Arg Pro Arg Arg Ile Leu Glu Val Gly Cys Gly Thr 2105 2110 2115	6354
ggg ctg atg ctc ttt gcc ctg ctg ccg cac tgc gag cgc tac gtg Gly Leu Met Leu Phe Ala Leu Leu Pro His Cys Glu Arg Tyr Val 2120 2125 2130	6399
ggc acg gat ttc tcc ccg gcg gcg ttg gac tac gtg cgt cgc tac Gly Thr Asp Phe Ser Pro Ala Ala Leu Asp Tyr Val Arg Arg Tyr 2140 2145 2150	6444

2135	2140	2145	
ctc ccc ccg gag cac ccg ggc cgc gtg gag ctg cta cac cgc acg Leu Pro Pro Glu His Pro Gly Arg Val Glu Leu Leu His Arg Thr 2150 2155 2160			6489
gcg gac gag tgg agc ggc gtg gcg gcc ggc tcc ttc gac gcg gtg Ala Asp Glu Trp Ser Gly Val Ala Ala Gly Ser Phe Asp Ala Val 2165 2170 2175			6534
ctg ctc aat tcg gtg gtg cag tac ttc ccc tcg cag gag tac ctg Leu Leu Asn Ser Val Val Gln Tyr Phe Pro Ser Gln Glu Tyr Leu 2180 2185 2190			6579
cgg cag gtg ctg gcc cgc tgc gtg gag gcc gtg gag gac ggg ggc Arg Gln Val Leu Ala Arg Cys Val Glu Ala Val Glu Asp Gly Gly 2195 2200 2205			6624
ttc gtc ttc gtg ggg gat gtg cgg agc ttg ccg ctg ctg gag tcc Phe Val Phe Val Gly Asp Val Arg Ser Leu Pro Leu Leu Glu Ser 2210 2215 2220			6669
ttc cac gcc tcg gtg gag ctg gag cgg gcc gcc ccc tcc atg ccg Phe His Ala Ser Val Glu Leu Glu Arg Ala Ala Pro Ser Met Pro 2225 2230 2235			6714
ctg gag gcg tgg cgc gag cgg gtg cgg cgt gcg gtg ctg gag gac Leu Glu Ala Trp Arg Glu Arg Val Arg Arg Ala Val Leu Glu Asp 2240 2245 2250			6759
aac gag ctg gtg gtg gac ccg gca ctc ttc gtg gcg ctg gcc cat Asn Glu Leu Val Val Asp Pro Ala Leu Phe Val Ala Leu Ala His 2255 2260 2265			6804
cag cat cca cgg gtg agc cac gtg gat atc gag ctg acg cgc ggc Gln His Pro Arg Val Ser His Val Asp Ile Glu Leu Thr Arg Gly 2270 2275 2280			6849
acg cac ccg aat gag atg gca cgc ttc cgc tac aac gcc gtg ctc Thr His Pro Asn Glu Met Ala Arg Phe Arg Tyr Asn Ala Val Leu 2285 2290 2295			6894
cac atc ggg ccg cgc acg ccg cca ccg gcc tcc gag gtg ccg tgg His Ile Gly Pro Arg Thr Pro Pro Pro Ala Ser Glu Val Pro Trp 2300 2305 2310			6939
gtg gac tgg tcc acg caa gga ctg agc ctc gat gca ctg cgg gct Val Asp Trp Ser Thr Gln Gly Leu Ser Leu Asp Ala Leu Arg Ala 2315 2320 2325			6984
cgg ctc cgg cag ggc ccg cca ggg cca ctg ggg gtg gca ggc att Arg Leu Arg Gln Gly Pro Pro Gly Pro Leu Gly Val Ala Gly Ile 2330 2335 2340			7029
ccc aat gcg cga gtg ctg ccc gcg gtc cgg gcc gcc gag gcg ctg Pro Asn Ala Arg Val Leu Pro Ala Val Arg Ala Ala Glu Ala Leu 2345 2350 2355			7074
ggc agc aca ggc agc gcg cgg cgg gtg gag gag ctg cgc cgg cgc Gly Ser Thr Gly Ser Ala Arg Arg Val Glu Glu Leu Arg Arg Arg 2360 2365 2370			7119
ctg tcc caa ccc ggc aca ggg gcc cag gat cca gac ccc ttc tgg			7164

239

Leu Ser	Gln Pro	Gly Thr	Gly	Ala Gln	Asp Pro	Asp	Pro Phe	Trp	
2375			2380			2385			
caa ctg	gcc gag	tcg ctc	ggc	tac acg	gcg gca	gtg	agc tgg	tct	7209
Gln Leu	Ala Glu	Ser Leu	Gly	Tyr Thr	Ala Ala	Val	Ser Trp	Ser	
2390			2395			2400			
cct ggg	cgg agg	gac ggg	gcc	ttc gac	gtg ctc	ttc	ctc ccg	gcg	7254
Pro Gly	Arg Arg	Asp Gly	Ala	Phe Asp	Val Leu	Phe	Leu Pro	Ala	
2405			2410			2415			
acg cca	ggg atg	cac ccg	cgg	tgg ctc	ggc ccc	acg	ccg ctc	aac	7299
Thr Pro	Gly Met	His Pro	Arg	Trp Leu	Gly Pro	Thr	Pro Leu	Asn	
2420			2425			2430			
cgc acg	ccg ccc	cca gcc	tcc	gcg cgt	ctg tcc	tcg	gag cct	cgg	7344
Arg Thr	Pro Pro	Pro Ala	Ser	Ala Arg	Leu Ser	Ser	Glu Pro	Arg	
2435			2440			2445			
cgg gcc	agc ctc	tca ctg	agg	ctg ggc	tcc gcg	ctg	cgg gcc	cac	7389
Arg Ala	Ser Leu	Ser Leu	Arg	Leu Gly	Ser Ala	Leu	Arg Ala	His	
2450			2455			2460			
ctg cag	acc cat	cta cca	gac	ttc atg	gtg ccc	tca	cgc ttc	gtg	7434
Leu Gln	Thr His	Leu Pro	Asp	Phe Met	Val Pro	Ser	Arg Phe	Val	
2465			2470			2475			
gtg ctg	cag tcc	ctg ccg	ctc	act ccc	aat ggc	aag	gtg gac	cgc	7479
Val Leu	Gln Ser	Leu Pro	Leu	Thr Pro	Asn Gly	Lys	Val Asp	Arg	
2480			2485			2490			
gcg gcc	ttg ccc	gtc ccc	gac	tcg cgt	cgg ctg	gaa	tct gcc	ccg	7524
Ala Ala	Leu Pro	Val Pro	Asp	Ser Arg	Arg Leu	Glu	Ser Ala	Pro	
2495			2500			2505			
ctg gtg	ccc ccc	agc aac	gag	ctg gag	cgg gtg	ctg	gcg cag	gtg	7569
Leu Val	Pro Pro	Ser Asn	Glu	Leu Glu	Arg Val	Leu	Ala Gln	Val	
2510			2515			2520			
tgg aag	gaa gtg	ctg gga	ctg	gag gaa	gtg agc	cgg	gag gac	aac	7614
Trp Lys	Glu Val	Leu Gly	Leu	Glu Glu	Val Ser	Arg	Glu Asp	Asn	
2525			2530			2535			
ttc ttc	gac gtg	gga ggg	cac	tcg ctg	ctg ctg	gcg	cag gtg	tgc	7659
Phe Phe	Asp Val	Gly Gly	His	Ser Leu	Leu Leu	Ala	Gln Val	Cys	
2540			2545			2550			
agc cgg	ctg gag	gcc cga	ctg	ggc cgg	cgg ctc	gaa	ctg gtg	acc	7704
Ser Arg	Leu Glu	Ala Arg	Leu	Gly Arg	Arg Leu	Glu	Leu Val	Thr	
2555			2560			2565			
ctc ttc	cgc tac	tcc tcc	att	gct gcc	ctg gcc	gag	cac ctg	cag	7749
Leu Phe	Arg Tyr	Ser Ser	Ile	Ala Ala	Leu Ala	Glu	His Leu	Gln	
2570			2575			2580			
gcc ccc	cag gag	ctg gca	gcc	gcc gag	gcg cag	gtg	cag cgc	atg	7794
Ala Pro	Gln Glu	Leu Ala	Ala	Ala Glu	Ala Gln	Val	Gln Arg	Met	
2585			2590			2595			
gcc acg	gag cgc	gcc ttg	ctc	cag caa	cag gca	gcc	caa cgc	cgc	7839
Ala Thr	Glu Arg	Ala Leu	Leu	Gln Gln	Gln Ala	Ala	Gln Arg	Arg	
2600			2605			2610			

accctggagt atgccctggc gcagacctgg ctgggctggg ggctccaacc gcaggcgctg 9748
gcaggccata gcctgggcca gctggtggcg gcgtgcctcg cgggtgtctt ctccctggaa 9808
gatgcacttc agctcgtggc agcccgggga cagctcatgc agggatgccc accggggggc 9868
atgctggccg tgccgctgcc ggaggccgag ctggcgggcg tgctgggcag cgagctgtgt 9928
atcgcccgcg tgaacggggc acgggcctgc gtggcctcgg gtctctctacc ggcggtagag 9988
gcgtcaccg ccgcgctgga gagccggggc gtgtccagcc ggcgcctgga gacttcccat 10048
gccttccact ctgcctccat ggaggcatgc cagggggccac tcaccacgt gctgcggcg 10108
atgcgcctgc agggccctcg cctcccctgt gtctccggcc tcaccggccg gtggctcacc 10168
ggcgaggagg ccacggagcc tacgtactgg gcccgcagc tccgcgagcc ggtacgtttc 10228
tccgaggccc tcgagacgct ctggagcctc aaggagccag tgctgctgga ggtaggcccc 10288
ggcaccacgc tcaccgcgt tgcgcgggcg caccacccc gcccgccccg caccaggag 10348
gtggcgagcc ttcccgctga accggacacg gccgtgccgt gcattgaaga ggccgtgggt 10408
gagctgtggc aggcgggcct ggagctggac tggagcgccc tgacgcgcgc gccacgccac 10468
cgcgcacacc ttccacccta ccccttcgag cgccagcgct actggattga accggaggct 10528
gccccgcagc cgcgcgcgca gcagccgacg cccgcgtccc ttgtgcccc ggagcagccc 10588
tcgcgcgagg ccctggaaga ctggttctac gtgccacact gggagcaggc acctgccacg 10648
agcggggggg gacagccct ggcgggcccc gtgctgcct tcatggactc ctcgggcctg 10708
gccgagcagg tgctggccgc gctgtggccc gcgactccg gcgcgctcct caccgcgtc 10768
gagccggccg ggcactatga gcaactcagt gagcacgcct tccgcctccg ccccgagagc 10828
gaggaggact gggacgcgt cttccaggcg ctccagtccc agggccgggt cccccgtcg 10888
atcctccatg cctgggctct caccgcggag cccggcccct gtaccccgga cggggaagcg 10948
gtgctggaac agggcttctt cagcctgctg cggctggccc gtgcgctcgg ccgccacgcg 11008
cctgagcgcc ctgtacagct cgaggtgctg tccagcttcg tccacgcggt gggccccgcg 11068
gagccgctcg agccgctcaa ggccacgctg ctgggggcct gtgcggtact tcccctggag 11128
taccctcacg tgcaagtgcg caccattgac gtgcggcccc ggagcgagcc acgggaagtg 11188
ctggtgagga gcctggccgc cgagctggcc gctcccatgg gagagagccc ggtggcctgg 11248
cgcgacgggc agcggtacgt gcgcgcgcgc acccggcaga ggctcgaggc ctgcggcccc 11308
ctgcgcagcc ttcgggagcg aggcgtgtac ctggtagcag gagggtggg tggaatcggc 11368
ctggtgctgg cgcgagcgct ggcccagcg gcgcgtgctc ggctggcgct actcaccac 11428
tcgcccttcc ctccccgaga gcagtgggag cagtggctgg aggaagcccc ggcgcacccg 11488
gagccggcgt ggcggagcga ggccgacccc tcggagcgcc gcaggacca gcaccgcac 11548

cgctgcctgc tggagctgga gcagctcggc gccgaggtgc aggtgtacac ggcggacgtg 11608
gcggaggagg ccgccgtgcg ctccggtggtg gagcaggtgc acgcccgtg ggggaagatc 11668
cacggcgtgc tgcacgccgc cgccacgttc gacgacggcg tcatccagct ccgcacgcac 11728
gagcagtcct cgcggggcct gcgcaccaag gtgcggggca gcatggtgct gcacgaggtg 11788
ctggcgagcg aggggttgga ttggttcgcc ctgtgctcct cgctggcgtc ggcgctgggc 11848
tcattcggcc aggcggacta ctgcggggcc aatgccttcc aggatgcgta tgcgcaccac 11908
ctgcgccggc agggcttcac gggagcgctg gcgctggact ggggcacgtg gagagatacg 11968
ggggcgggca tgcggctggt ggcacgcacc cgacgggggg gccatgagaa gccgcccacg 12028
ccgctcacc accactctt cgattgcgaa cagcgcgagc cgggcgggac gcactggctc 12088
ggcctcacgc tgaggggtgg agaggactgg gtggtggacg agcaccggct gcaaggggtg 12148
ccaacactgc ccggggtggc ctacctggag ctggcgcggg cggcggtgtc ccaggcgctg 12208
ggggccgagg cggtaggact ggcggagctg ttgctgctgg agcctctgac ggtacctga 12268
ggcgaatcca ggcaggtccg ggtggtgctc cagcccagag ggcaggctca tgccctgcgg 12328
gtggagagcc ggtcggagga agcgcgggga tggaatgagc atgcgcgggg ccgggtgcgc 12388
gccgtgcctc ggctggccga gcgcattccag cccgagctgc tgcgcgccgc ctgcgaacac 12448
gagcagcccg tgcccggaga gccccaggag caaggccctg tccacgccgg agcacgctgg 12508
catggcctct tccagtgggt tcgccggggc cctcgccagg cctcgccca gctcgcgcta 12568
ccggagccct tccacggtga cctcgagcgc ttcgagctcc acccggcgt catggacatg 12628
gccaccagct tcgccattcc cggcggagtg ccctggctcg cctttggcta tgagcgcggtg 12688
ctcattcacg gtccgctgcc accgcaggtg ctacagccag tgagcctgcc cgaggagtcg 12748
caggctggcg cgcaacaact ccggctacag gtgcggctgc tggatctgga gggctgggag 12808
cgggtgcgga tcgatgggta cctgctgcgg ccgctgaagc ccagcgacgc cagtgtcgag 12868
ccggcgggcg ccaacgtgga agtcgccgtg gggacgccgg ggctgctgga gagccttggc 12928
ctgcgccgct gcacccgccc cgccccggg ccgcgccagg tggagatcga ggtggaggcc 12988
gctgggctca acttcctgga cgtgctgggc gcgctgggga tgatgccggc attggaggcg 13048
gaggagagcg tactggggcg cgagtgtcc ggacgcattg ccgccgtggg cgagggcgtc 13108
agcgggctgc gcgtggggga cgaggtgctg gcggtggccc caggctgctt ccgctcctac 13168
gtgctgggtg atgagagcca ggtggtgcgc aggcccgct cgctggggct cgccgagggg 13228
cccccaccga tggtgccgtt cgccacggcg tacttcgccc tgcacaccgt gggccgggtg 13288
ccggcgcccg agcgcattct catccacgcc gcggccggag ggctgggtct ggccgccgtt 13348
cagctcgcat ccgggaccgg ggcggagata ttggccaccg cgggcagtga gcagaagcgc 13408
gagtacctgc gctcgctcgg cattgctcac gtgctggact cgcgcagcac ctcttcgctc 13468

agcgaagtgc gcgagcgcac cggcgggctg ggggtggatg tgggtctcaa ctgctggcg 13528
ggagagctgc tcctggcggg cctgtccgtg ctggccccgc acggccgctt cctggagctg 13588
gggaagaggg acctgtatgc ggaccagcag gtgggcctcc gaaccctggc ccgagggcag 13648
actttcgccg ccattgactt cgggtcccccac caccgggact tccgagcggg gctcgaggag 13708
gtggccacgc aactcaccca gggccagctc gagccattgc ccaccgcct cttccccgcg 13768
cggcaggtgg ccgaagcctt cagcttcatg gcccgcgcg cgcacatcgg tcgcgtcgcc 13828
gtctccatgc agggggcgac ggcattgccc gcgtccatga ctcggggctc caggcccgcg 13888
ccggtggccg tacctccttg ggaggacccg cggctggcgg gcggcatctc ctctgaagag 13948
ggtgcggagg ccttcctgcg cgcgttgag caaggggcac cgcagctcat catctcccc 14008
caggacttca gctcgtgct gcgcggcctg ggcggcagcc agggcgctgcg cgaaaaggag 14068
cgctcgtca ccgggcgcgc cgcgcgcgc gagccgcagg ccttgccacc ctctcgtg 14128
gagcagctca tcgagcagg gtggcgcaag cacctgggtg tggagcgcgt gcagcccacg 14188
gacagcttct tcagcttg aggagactcg ctgctgggca tccaagtggc ggcggatctc 14248
cgcaggcacc tgggtgtgga gctgcccac gccaccctct tcagccacc caccctcgcc 14308
gcgctggccg cggctctgcg ggctcgacag ggtgaggccg cggctccac tgccccgcg 14368
ccagcgctcg tgccggatcc agccgcgcgc ttcgagcctt tccccctcac ggatgtgcag 14428
gaagcctact ggggtgggtcg ccgctcggcc ttcgagctcg gcggcgctcg cgcctatggc 14488
tacttcgaaa tcgaaagccc ggggctggag gtggagcgct tcatccaatg ctggcgccag 14548
ctcctgcagc gccacgacat gctgcgcatg gtggtgctcc cggatgggcg gcagcagggtg 14608
ctcgagcagg tgccggagta cacgccggag gtggtggagc tgccggggct cttccccag 14668
gaggccgagt cccgccggct ccagctcgt gagcgcatgg cccaccagg cctgcgcagc 14728
gaccgctggc cgctcttcga gctggtgctc tgccggtagc agggaggcgt ccgcatccac 14788
atgagcatgg atgccctgat gctggatgcg tggagtccag cgggtgcttcg gcaggacttc 14848
gcccagctgt accacgagcc gggccggccg ctggagccgc tggccatcac cttccgcgac 14908
tacgtgctgg cggagcgccg gctgcgcgag ggcgaggccc atgagcgcg ccgcgcatac 14968
tggtgggctc ggctggacac gctgccgcca ccgccgagc tgcccctggt gaaggaaccc 15028
tcgcagctgg agcacgcgc gttcacccac cgcgaggctc ggctcgagcc acaccgctgg 15088
gcccggctcc aggagcgggc gcgcgcccac ggcctcacc cctcgggcg ctgcatggcc 15148
gccttcgccc aggtgcttgc ccgctggagc cgtcaccgc gcttcaccct caacctcacc 15208
ctcttcagc gcttgccct gcacccgcag gtggagcagc tgggtggcg cttcacctcc 15268
ctggtcctgc tggaggtaga ggcacacgc gcgagcacct tcgccgagc tgccctcccg 15328

ctccaggcac agctatggcg ggacctggag cacggcagcg tgagcgccgt gcagctcatc 15388
cgcgagctcg tccgcaccgg ccgccgtcc ccgggcgcca tcatgcccgt cgtcttcacc 15448
agcatcctca gcctggatgc gcggcgcggc cccaggggca gcctctcctt cttcgaggga 15508
gaactggtgt acagcatcag ccagactccc caggtgtggc tggaccacgg agtccacgag 15568
gaggaggggg cgctcgtcct ggcgtgggac tcggtggagg cgctcttccc tccgggcatg 15628
gtggacgaca tgttccacgc ctaccagcgg ctgctggggg cactcgccga ggaggagcaa 15688
gcgtgggagg gcgagctgcc ggagctgctg cctcctgccc agcgtgagtt gctcgcgcgc 15748
tacaacgcca cccaggcgcc gcggcccagc ggacggctgg aggagggtt cttcaccag 15808
gcggggctcc accccgagct ccccgcgctg ctgcgaccgg agcgcaccct gagctacggc 15868
gagctggcaa ggcgagccca ggcgctggcc gcgcgcctac gcgagctgga ggtgcagcct 15928
caggagttagg tggccattgc catgcacaag ggctgggagc aggccaccgc cgtgctcggc 15988
gtgctgcagg cggccgcggc ctacctgccg ttggatccag agcagccgcc gctccggctc 16048
caccagctcc tggaggagg gcccgcgcgc gtggtgctca cccagtcctc cttgctgcac 16108
accgtgccct ggccgccggg tgtgcaggtg atcgccgtgg acgaactcga gcctgcaacc 16168
gaggcccccc cgctaccgcc tcgcggtaca ccggagcacc tcgcctacgt catctatacc 16228
tccggctcca ccggcaaacc caagggtgtg gccattgaac accgggccgc gctcaacacc 16288
gtggtggacc tcaacacccg ctttggcgtc ggcccggagg atcgggtgct gggcctgtcc 16348
gcgctcacct tcgacctgtc ggtgtacgac gtgctggggc tgctcggcgc gggcggagcg 16408
ctggtgctgc ccgcggcgga ggcggagaag gatccccgcc actggtggga gcggctggtg 16468
gctggggcggt tcacggtgtg gaactccacc ccggcgctga tgctgctgct ggtggagtac 16528
gccgagcagc gcgggctgaa gctgcccgcg gcgctgcgcc tggatcatgct cagcggcgac 16588
tggatccccg tggcgctacc ggatcgcatc cgcgactcg gcagggacgt gcaggtggtg 16648
agcctgggag gcgccaccga ggcttccatc tggctcattg cctaccccat tggccagggtg 16708
gcaccgcagt ggaagagcat tccctatggc atgccgtgg ccaatcagcg attccacgtg 16768
ctggacgggc ggctggaggc acggccctgg tgggtgccgg gcgagctgta cattggcggc 16828
gaggggctgg cacgcgagta ctggagggat gagccgctca ccgcaacgcg cttcatccgt 16888
caccgcgcga caggcgagcg gctgtaccgc accggggacc aaggcgcat gctgcccga 16948
gggagcatcg agttcctcgg ccgggaggat ctgcaggtca aggtgcaggg cttccgcgtg 17008
gagctggggg agattgaagc ggcgctggcc cagcaccgg cgctatcggc gagcgtggtg 17068
gtggcacgag gagagcctcg gggagtgcgg cggctggtgg cctacgcggt tcctcgctcg 17128
ggtcagacgc ccgcggcagg cgagctgcgc cgctacctgg cggagcggct accgcgtac 17188
atggtgccct ccgccttcgt cctcctggag tccctgccc gcagccgcaa cgggaagatc 17248

gcccagagatc agctccccga gccgcagcag acacagggac tggctgcaca ggccgcggcg 17308
 gccgacccgc tcgtggagcg gctggcagcc ctggtgaaag aggcgctgag gctcgagcgc 17368
 gtggagcccc aggacagcct gctggacctg ggggcggact cgggtggcgct catccgcctc 17428
 atcaaccgcc tggaggcaga gctgcagttc cgcccgcgac tggccgacat ctacgagaac 17488
 cccaccgtgc agggggctcgc cacgctccac caagagaaga caaagagcca gggagagggga 17548
 ggcgctccgc ggctcacggc gcccgcctcc acgcttctgc ccgccgagga gtggggggcg 17608
 ttcaaggcca accgcccggg cctgcgcgcg tccccgatg gaacgcctga ggtggcactg 17668
 ccaggcagcg gcttggcgcc ggcccccgag gagctgacag ctctcgagcg ccgccgcagt 17728
 gtgcgcacct actccctcga gcccgtagac catgagcagt tgggacggct gctggccccc 17788
 ttgcgcgagt gggaggtgca gggctcgcgg cgttacctct acgcctccgc gggcgggctc 17848
 tacccggtgc agctctacct ccacctcaag cccggccggg cgcgagggct cgagcccggc 17908
 acctggtatt acgatcccag caccatcgg ctggtgctcc tgtccgccgg agccgggctg 17968
 gatcgccgca tccatgatcc gcaccagaac caggccatct tcgactcggc cgcgttctcc 18028
 ctcttcctca tcgcccgcac gggggccgtg gagccggtat acgccgaaca tgccctgcac 18088
 ttcgccacgc tggaggcagg gctgatgacc cagcttctcg atctgggcgc ggctcccagt 18148
 gggctggggc tgtgccatat cggcgacctg gacttcgccc aggcgcgggg gctcttcac 18208
 ttggaggagg agcacgtgct gctgcacagc ctggtggcg gagtactccc cactcggggg 18268
 caggaggcgg cctcggtgcc cgctgagggg ggaacggagg cgcggaact ggcgcagctc 18328
 ctgcaacagg tgaagacgct cacgcccag gccgcgcgag cgctgctcga ggcccgcgcg 18388
 ggcagcaagg ggagaccac atg agt aag ctg cac gag gag ctg gag agc 18437
 Met Ser Lys Leu His Glu Glu Leu Glu Ser 2635
 ctg gcg ccg gag cag cgg gag cta ctc gcg gcc ctg atg aaa gag 18482
 Leu Ala Pro Glu Gln Arg Glu Leu Leu Ala Ala Leu Met Lys Glu 2650
 cag ggc ctg gac gag ggc gca ttg ctg atg ccc gtg gag cgc aag 18527
 Gln Gly Leu Asp Glu Gly Ala Leu Leu Met Pro Val Glu Arg Lys 2665
 ccc gag ggg ctg ccg ctc tcg tcc gcc cag cag cgc atg tgg ttc 18572
 Pro Glu Gly Leu Pro Leu Ser Ser Ala Gln Gln Arg Met Trp Phe 2680
 ctg caa cag ctc agg cct acg agc ccc ttc tac aac gtg cac gcg 18617
 Leu Gln Gln Leu Arg Pro Thr Ser Pro Phe Tyr Asn Val His Ala 2695
 gcg ctc cgg ctc acg gga aag ctg aag gtg gag tgc ctg gtg cac 18662
 Ala Leu Arg Leu Thr Gly Lys Leu Lys Val Glu Cys Leu Val His 2710

246

agc ctg aat gag ttc	gtc cgg cgt cac gag	ccg ctg cgc acg gtg	18707
Ser Leu Asn Glu Phe	Val Arg Arg His Glu	Pro Leu Arg Thr Val	
2715	2720	2725	
ttc ccc tcc gcc gga	ggg caa ccc ctg cag	cgc atc ctc gcg ccc	18752
Phe Pro Ser Ala Gly	Gly Gln Pro Leu Gln	Arg Ile Leu Ala Pro	
2730	2735	2740	
gca ccg gct gcg ctc	gag cag cgc gac ctg	agc ggc gtg ccc gcc	18797
Ala Pro Ala Ala Leu	Glu Gln Arg Asp Leu	Ser Gly Val Pro Ala	
2745	2750	2755	
cag gag cga gag gcc	gag gtg tac cgc gcg	gtg gag cac gca gcg	18842
Gln Glu Arg Glu Ala	Glu Val Tyr Arg Ala	Val Glu His Ala Ala	
2760	2765	2770	
ctc gcc tcc ttt gat	ctg gag cga gag ccg	ccg tgc cgc ttc ctc	18887
Leu Ala Ser Phe Asp	Leu Glu Arg Glu Pro	Pro Cys Arg Phe Leu	
2775	2780	2785	
ctg gtg cga gtg gag	ccg gag gaa cac gtg	ctg gtg ttc gcc acg	18932
Leu Val Arg Val Glu	Pro Glu Glu His Val	Leu Val Phe Ala Thr	
2790	2795	2800	
cac cac atc gca gcg	gat ggc tgg tcg ctc	ggc gtc ttc gtg cag	18977
His His Ile Ala Ala	Asp Gly Trp Ser Leu	Gly Val Phe Val Gln	
2805	2810	2815	
gag ctg tgc gcg ctg	tac acg gct gct gta	cac gcg gag cca ccg	19022
Glu Leu Cys Ala Leu	Tyr Thr Ala Ala Val	His Ala Glu Pro Pro	
2820	2825	2830	
gcg ctg cca ccg ctg	cgg ctg cag tac gcg	gac ttc gcc gcg tgg	19067
Ala Leu Pro Pro Leu	Arg Leu Gln Tyr Ala	Asp Phe Ala Ala Trp	
2835	2840	2845	
gag cgg agc cgg ctc	aag ggt ggg cgc gag	cgc gag ctg ctc gag	19112
Glu Arg Ser Arg Leu	Lys Gly Gly Arg Glu	Arg Glu Leu Leu Glu	
2850	2855	2860	
tac tgg cag gag cag	ctg gca ggg ctg ccg	gac ctg agc acc ctg	19157
Tyr Trp Gln Glu Gln	Leu Ala Gly Leu Pro	Asp Leu Ser Thr Leu	
2865	2870	2875	
ccg cca gag cgg ccc	agg ccc ccg ctg tcc	aag cag gag ggc gcc	19202
Pro Pro Glu Arg Pro	Arg Pro Pro Leu Ser	Lys Gln Glu Gly Ala	
2880	2885	2890	
acg ttc gaa ttc gcc	ctg cca cca cac cag	gtg cag gcc ctg cgc	19247
Thr Phe Glu Phe Ala	Leu Pro Pro His Gln	Val Gln Ala Leu Arg	
2895	2900	2905	
tca ctt gct caa gcg	cgg cgc acc tcg ctc	ttc tcc gtg ctg ctc	19292
Ser Leu Ala Gln Ala	Arg Arg Thr Ser Leu	Phe Ser Val Leu Leu	
2910	2915	2920	
gcc gcc ttc caa tgg	ctg ctg gcc cgg tgc	gcc ggc cag gac gac	19337
Ala Ala Phe Gln Trp	Leu Leu Ala Arg Cys	Ala Gly Gln Asp Asp	
2925	2930	2935	
gtg gcc ctg ggc atg	ccc atc gcc aac cgc	aac cgc aag gag ctg	19382
Val Ala Leu Gly Met	Pro Ile Ala Asn Arg	Asn Arg Lys Glu Leu	
2940	2945	2950	

gaa ggc ctc atc ggc	tgc ttc gcc agt acc	ctg gtg ctg cgc gcg	19427
Glu Gly Leu Ile Gly	Cys Phe Ala Ser Thr	Leu Val Leu Arg Ala	
2955	2960	2965	
aag ccc tcc gcc tcg	acc gcc ttc acc tca	tgg ctc gcc cag gtg	19472
Lys Pro Ser Ala Ser	Thr Ala Phe Thr Ser	Trp Leu Ala Gln Val	
2970	2975	2980	
tcc gag cag ctc cac	ggc gcc ctc gag cac	cag gag gtg cct ttc	19517
Ser Glu Gln Leu His	Gly Ala Leu Glu His	Gln Glu Val Pro Phe	
2985	2990	2995	
gaa cga ttg gtg gag	gtg ctc cag ccc cgg	cgc cgc atg gat cgg	19562
Glu Arg Leu Val Glu	Val Leu Gln Pro Arg	Arg Arg Met Asp Arg	
3000	3005	3010	
cac ccg ctc ttc cag	atc ttc ctc gcc atg	cag cag cac ccg ctg	19607
His Pro Leu Phe Gln	Ile Phe Leu Ala Met	Gln Gln His Pro Leu	
3015	3020	3025	
cgg cgc gcc gag ctg	cca ggc ctg ctg ctc	tcg gag ttc ccc ctt	19652
Arg Arg Ala Glu Leu	Pro Gly Leu Leu Leu	Ser Glu Phe Pro Leu	
3030	3035	3040	
cgc agc cgc gtg gcc	cgc ttc gat ctg gag	ttc cac ctc tgg gag	19697
Arg Ser Arg Val Ala	Arg Phe Asp Leu Glu	Phe His Leu Trp Glu	
3045	3050	3055	
tct gcc cag ggc gtc	gag ggc acc ctc atc	tac gac gtg gac ctg	19742
Ser Ala Gln Gly Val	Glu Gly Thr Leu Ile	Tyr Asp Val Asp Leu	
3060	3065	3070	
tat ggc cag gag ggc	gta gcg caa ctg gtg	cgg cgc tac gtg agc	19787
Tyr Gly Gln Glu Gly	Val Ala Gln Leu Val	Arg Arg Tyr Val Ser	
3075	3080	3085	
ctg ctc gaa gcc gtg	gcg gcg gac ccg acg	cgg aca ctg ggg gag	19832
Leu Leu Glu Ala Val	Ala Ala Asp Pro Thr	Arg Thr Leu Gly Glu	
3090	3095	3100	
ctg gcg ggt gag acg	ctg gcc ccg ccc gta	cgg aag cag gtt ctc	19877
Leu Ala Gly Glu Thr	Leu Ala Pro Pro Val	Arg Lys Gln Val Leu	
3105	3110	3115	
gcc ctg tcg gag tgt	ccg ccc ctc cca cgc	ccg agc aca gcc cct	19922
Ala Leu Ser Glu Cys	Pro Pro Leu Pro Arg	Pro Ser Thr Ala Pro	
3120	3125	3130	
cgg aca ctg gcc gag	gca ctt cta cag acc	gcc gag cgc ttc ccc	19967
Arg Thr Leu Ala Glu	Ala Leu Leu Gln Thr	Ala Glu Arg Phe Pro	
3135	3140	3145	
acg gcc acg gtc agc	ttc gtc cag gca cag	ggc tct tgc acg gcc	20012
Thr Ala Thr Val Ser	Phe Val Gln Ala Gln	Gly Ser Cys Thr Ala	
3150	3155	3160	
tgg act ctg cca gag	ctg gtg gag cgg gcc	cgg cgc ctg cag gcc	20057
Trp Thr Leu Pro Glu	Leu Val Glu Arg Ala	Arg Arg Leu Gln Ala	
3165	3170	3175	
ggc ttg cgg cag tgg	gga ttg cgg ccg gcc	gac agt ctg gtg ctg	20102
Gly Leu Arg Gln Trp	Gly Leu Arg Pro Gly	Asp Ser Leu Val Leu	

248

3180	3185	3190	
gtg ctg ggg agg gag Val Leu Gly Arg Glu 3195	gaa gag acg gta gag Glu Glu Thr Val Glu 3200	gca ctc tgg gcc tgc Ala Leu Trp Ala Cys 3205	20147
gtg ctg gcg gga gtg Val Leu Ala Gly Val 3210	gcg ccc ctg gtg cta Ala Pro Leu Val Leu 3215	ccg gcg ccg ccc gcg Pro Ala Pro Pro Ala 3220	20192
aga gcg gag gcg agt Arg Ala Glu Ala Ser 3225	ccc gcg ctc tca cgt Pro Ala Leu Ser Arg 3230	ctg cgg cat gcc ccg Leu Arg His Ala Arg 3235	20237
cag ctg ctg ggt ggg Gln Leu Leu Gly Gly 3240	ccc ccg gtc ctc acc Pro Arg Val Leu Thr 3245	ccg cag gag atg ctg Arg Gln Glu Met Leu 3250	20282
ccc gat ctt gcc agg Pro Asp Leu Ala Arg 3255	cag ctt cag gtc tcg Gln Leu Gln Val Ser 3260	ccc acc gcc gac atc Pro Thr Ala Asp Ile 3265	20327
ctg ggg gcc gtg gag Leu Gly Ala Val Glu 3270	gag ctg aga gcc acg Glu Leu Arg Ala Thr 3275	ggt ggc gag gcg ccc Gly Gly Glu Ala Pro 3280	20372
ctg ccc cca ggg cgc Leu Pro Pro Gly Arg 3285	atg gat gac gtg gcc Met Asp Asp Val Ala 3290	ctg ctc aac ctc acc Leu Leu Asn Leu Thr 3295	20417
tcg ggc acc acc ggc Ser Gly Thr Thr Gly 3300	aag gcc aag tgc gcc Lys Ala Lys Cys Ala 3305	atg ctc acg cac cgc Met Leu Thr His Arg 3310	20462
aac ctg ctg gtg cgc Asn Leu Leu Val Arg 3315	ctg gag gcc acc aac Leu Glu Ala Thr Asn 3320	gtc gtc tac gag tcc Val Val Tyr Glu Ser 3325	20507
cag ccg ctc gag cgc Gln Pro Leu Glu Arg 3330	ggc ctc gtc tgg ttg Gly Leu Val Trp Leu 3335	cag ctg cac aac att Gln Leu His Asn Ile 3340	20552
ggc gcc ctg tcc gag Gly Ala Leu Ser Glu 3345	tac cac ctc cgc ccg Tyr His Leu Arg Pro 3350	ctc tgc gcc ggc atg Leu Cys Ala Gly Met 3355	20597
cac acc ttc cac gct His Thr Phe His Ala 3360	ccc acc gag gag gtg Pro Thr Glu Glu Val 3365	ctc gcc gag ccc ctg Leu Ala Glu Pro Leu 3370	20642
cga tgg ctg gag tgg Arg Trp Leu Glu Trp 3375	ctc gag ccg tat ggc Leu Glu Arg Tyr Gly 3380	att gca cag acg tgg Ile Ala Gln Thr Trp 3385	20687
gcg ccc agc ttc gcc Ala Pro Ser Phe Ala 3390	tac agc cac ctg ctg Tyr Ser His Leu Leu 3395	gag ccg ctg cgc aag Glu Arg Leu Arg Lys 3400	20732
trg gag gac ccg cgc val Glu Asp Arg Arg 3405	tgg agc ctg gga ggc Trp Ser Leu Gly Gly 3410	gtg cgc gtg ctg ctc Val Arg Val Leu Leu 3415	20777
agc gcc gga gag cag	atc tcc gcg ccc atg	gtg gag gag ctg atg	20822

249

Ser Ala Gly Glu Gln	Ile Ser Ala Pro Met	Val Glu Glu Leu Met	
3420	3425	3430	
cgg agg ttg gct ccc	tcg gga gtg cgg gag	gat gcc ttc gtc gct	20867
Arg Arg Leu Ala Pro	Ser Gly Val Arg Glu	Asp Ala Phe Val Ala	
3435	3440	3445	
gcc tgg gga atg acc	gag acg gct tcg ggg	gtg acc tat gcg cgg	20912
Ala Trp Gly Met Thr	Glu Thr Ala Ser Gly	Val Thr Tyr Ala Arg	
3450	3455	3460	
cgg ccc gga acg ccg	ccc cgc atg cac acc	ctg gag cgc gcg agc	20957
Arg Pro Gly Thr Pro	Pro Arg Met His Thr	Leu Glu Arg Ala Ser	
3465	3470	3475	
ttg agc ggt ccg ctt	cgg cac gcg gca ccc	gct tcg ccc acc gcg	21002
Leu Ser Gly Pro Leu	Arg His Ala Ala Pro	Ala Ser Pro Thr Ala	
3480	3485	3490	
ctg cgg ctg atg gat	gta gga gcc ccc att	gcc ggc acc gcc ctg	21047
Leu Arg Leu Met Asp	Val Gly Ala Pro Ile	Ala Gly Thr Ala Leu	
3495	3500	3505	
cgg gtg gtg gac gcg	agc ggc gag ctg ctc	tcc gag gag tgc gtg	21092
Arg Val Val Asp Ala	Ser Gly Glu Leu Leu	Ser Glu Glu Cys Val	
3510	3515	3520	
ggg cgc att caa gtg	cgc ggc gag atg atc	tct ccc ggc tac tac	21137
Gly Arg Ile Gln Val	Arg Gly Glu Met Ile	Ser Pro Gly Tyr Tyr	
3525	3530	3535	
ggg gac ccg aag gcc	tcg gcg gcg ctg ctc	acc gcc gac ggc tgg	21182
Gly Asp Pro Lys Ala	Ser Ala Ala Leu Leu	Thr Ala Asp Gly Trp	
3540	3545	3550	
ctg gag aca ggc gat	ctg ggc ttc ctc tcc	gag gga gcg ctg acc	21227
Leu Glu Thr Gly Asp	Leu Gly Phe Leu Ser	Glu Gly Ala Leu Thr	
3555	3560	3565	
att acc ggc cgc gcc	aag gat ctg gtc atc	atc cac ggc acc aac	21272
Ile Thr Gly Arg Ala	Lys Asp Leu Val Ile	Ile His Gly Thr Asn	
3570	3575	3580	
ttc tcc tgc tac gag	atc gag tcc gcg gtg	gag cag gtg gag gga	21317
Phe Ser Cys Tyr Glu	Ile Glu Ser Ala Val	Glu Gln Val Glu Gly	
3585	3590	3595	
gtg gcg ccc tcc tcg	gcg gcg gcg gcg gcg	gtg cgg atg ctc gag	21362
Val Ala Pro Ser Ser	Ala Ala Ala Ala Ala	Val Arg Met Leu Glu	
3600	3605	3610	
ggg agc agg gag gag	ctg gcg gtc ttc ttc	gtg ccc acg gag gga	21407
Gly Ser Arg Glu Glu	Leu Ala Val Phe Phe	Val Pro Thr Glu Gly	
3615	3620	3625	
ctc gct ccg caa ccc	cct gcc tca ctc ctg	tcg cgc atc cgc cag	21452
Leu Ala Pro Gln Pro	Pro Ala Ser Leu Leu	Ser Arg Ile Arg Gln	
3630	3635	3640	
cag gtg ctc gag cag	gtg ggg gtg cgg atc	gat cat ctc atc ccg	21497
Gln Val Leu Glu Gln	Val Gly Val Arg Ile	Asp His Leu Ile Pro	
3645	3650	3655	

250

ctc gag cct cac cag	ctc cca cgc acg gag	gga ggc aag ctg cgc	21542
Leu Glu Pro His Gln	Leu Pro Arg Thr Glu	Gly Gly Lys Leu Arg	
3660	3665	3670	
cgc tcg gag ctg agg	gcg cgc ttc gaa gca	ggt gag ctg agg gcg	21587
Arg Ser Glu Leu Arg	Ala Arg Phe Glu Ala	Gly Glu Leu Arg Ala	
3675	3680	3685	
ccg caa ccc gct ccg	gtc ccg agc ccc tcg	cgc ccg ctg gag cag	21632
Pro Gln Pro Ala Pro	Val Pro Ser Pro Ser	Arg Pro Leu Glu Gln	
3690	3695	3700	
ctc atc gcc tcc gtc	tgg gcg gag gtg ctc	gag cac cag gac atc	21677
Leu Ile Ala Ser Val	Trp Ala Glu Val Leu	Glu His Gln Asp Ile	
3705	3710	3715	
gcg ccc gag gcc agc	ttc ttc gac ctg ggc	ggc aac tcc atc ctc	21722
Ala Pro Glu Ala Ser	Phe Phe Asp Leu Gly	Gly Asn Ser Ile Leu	
3720	3725	3730	
ctg gtg cgc gtg gag	cgg gca ctg cgc gcg	cgg ctc ggc ctg gag	21767
Leu Val Arg Val Glu	Arg Ala Leu Arg Ala	Arg Leu Gly Leu Glu	
3735	3740	3745	
ctg acg ctc atg gac	ctg ttc gcc tac ccc	acg gtg cac tcg ctg	21812
Leu Thr Leu Met Asp	Leu Phe Ala Tyr Pro	Thr Val His Ser Leu	
3750	3755	3760	
gcg gac tac ctg gag	cct cga gcc gcc cag	ctg cca gca cag gcc	21857
Ala Asp Tyr Leu Glu	Pro Arg Ala Ala Gln	Leu Pro Ala Gln Ala	
3765	3770	3775	
tcc cct ccc acg cag	gcc gag cgc cgg cgg	ggg atg cgc ggc gcg	21902
Ser Pro Pro Thr Gln	Ala Glu Arg Arg Arg	Gly Met Arg Gly Ala	
3780	3785	3790	
gcg ctg gag cag aga	cgc act cgc cgc cag	gca caa cgg gac gag	21947
Ala Leu Glu Gln Arg	Arg Thr Arg Arg Gln	Ala Gln Arg Asp Glu	
3795	3800	3805	
gac tag			21953
Asp			

<210> 34

<211> 2625

<212> PRT

<213> Angiococcus disciformis

<400> 34

Leu Ser Ala Gly	Ala Leu Leu Ala His	Ala Ala Ser Leu Gly	Val Arg
1	5	10	15

Leu Thr Val	Glu Gly Glu Arg	Leu Arg Phe Gln	Ala Pro Pro Gly	Val
20	25	30		

Met Thr Pro Glu Leu Gln Ser Arg Leu Gly Gly Ala Arg His Glu Leu
 35 40 45

Ile Ala Leu Leu Arg Gln Leu Gln Pro Ser Ser Gln Gly Gly Ser Leu
 50 55 60

Leu Ala Pro Val Ala Arg Asn Gly Arg Leu Ala Leu Ser Phe Ala Gln
 65 70 75 80

Gln Arg Leu Trp Phe Gln Glu Gln Leu His Pro Glu Ala Pro Ala Asn
 85 90 95

Asn Leu Thr Gly Ala Val Val Phe Thr Gly Pro Leu His Val Ala Ala
 100 105 110

Leu Leu Gly Ala Val Ala Ala Leu Val Arg Arg His Glu Ala Leu Arg
 115 120 125

Thr Thr Leu Gly Glu Glu Gly Gly Val Pro Tyr Ser Leu Ile Gly Glu
 130 135 140

Pro Trp Gln Pro Ala Leu Glu Val Glu Ala Leu Pro Gly Ala Thr Val
 145 150 155 160

Gly Glu Arg Leu Glu Gln Ala Arg Glu Val Ala Leu Ala Glu Ser Arg
 165 170 175

Arg Arg Phe Ala Leu Glu Thr Glu Pro His Leu Arg Val Arg Leu Leu
 180 185 190

Arg Leu Ala Glu Gln Gln His Val Leu Val Leu Ser Leu His His Ile
 195 200 205

Ala Ala Asp Gly Val Gly Leu Gln Val Leu Glu Gln Glu Leu Ala Ala
 210 215 220

Leu Tyr Gly Ala Leu Ser Ala Gly Ala Glu Pro Arg Leu Pro Pro Leu
 225 230 235 240

Pro Leu Gln Val Ala Asp Leu Ala Asp Trp Gln Arg Arg Trp Val Glu
 245 250 255

Gly Glu Glu Tyr Gln Val Gln Leu Ala Tyr Trp Arg Arg Gln Leu Ala
 260 265 270

Gly Leu Thr Pro Leu Glu Val Pro Gly Asp His Pro Arg Pro Arg Ile

275

280

285

Pro Ser Met Arg Gly Ala Glu Val Arg Ala Pro Leu Leu Ser Ala Pro
290 295 300

Gln Ala Gln Val Leu Arg Ala Leu Gly Gln Gly Glu Gly Ala Thr Leu
305 310 315 320

Tyr Met Thr Leu Leu Ala Ala Leu Gly Val Leu Leu Gln Arg Trp Thr
325 330 335

Gly Gln His Asp Met Ala Val Gly Ser Ala Ala Ala Asn Arg Asn Arg
340 345 350

Pro Gly Leu Glu Gly Ile Leu Gly Phe Leu Leu Asn Ile Val Leu Leu
355 360 365

Arg Leu Asp Leu Arg Gly Arg Pro Arg Phe Arg Glu Leu Leu Arg Gln
370 375 380

Ala Arg Arg Val Cys Val Glu Ala Tyr Ala His Gln Glu Leu Pro Phe
385 390 395 400

Glu His Leu Val Glu Ala Leu Gln Pro Gly Ser Glu Arg Gly Asp Ser
405 410 415

Ser Leu Tyr Arg Val Ala Leu Ala Val Ser Asp Thr Pro Trp Met Pro
420 425 430

Gly His Gly Leu Lys Leu Glu Gly Val Gln Ala Gln Pro Leu Asp Phe
435 440 445

Pro Arg Gly Val Leu Asp Leu Asp Leu His Leu Trp Val Tyr Asp Thr
450 455 460

Gly Glu Gly Leu Thr Gly Arg Leu Glu Tyr Ala Val Asp Leu Tyr Glu
465 470 475 480

Glu Pro Thr Ala Arg Arg Leu Leu Glu Gly Phe Arg Gln Val Leu Glu
485 490 495

Ala Val Val Glu Ala Pro Asp Arg Pro Val Pro Glu Leu Pro Val Leu
500 505 510

Gly Glu Gln Glu Arg His Gln Val Leu Ser Gly Trp Asn Arg Thr Gln
515 520 525

253

Arg Pro Tyr Pro Arg Glu Ala Ser Val His Gly Leu Phe Gln Gln Arg
530 535 540

Ala Leu Gln Ala Pro Arg Ala Val Ala Val Val Tyr Gly Glu Arg Ser
545 550 555 560

Leu Thr Tyr Gly Glu Leu Ala Glu Arg Ala Arg Gly Leu Ala Gln Gly
565 570 575

Leu Val Ala Arg Gly Val Arg Arg Gly Asp Leu Val Ala Leu Arg Leu
580 585 590

Glu Arg Ser Pro Glu Gln Val Glu Ser Met Leu Ala Val Leu Gln Ala
595 600 605

Gly Ala Ala Tyr Val Pro Leu Asp Pro Ser Tyr Pro Val Gln Arg Gln
610 615 620

Glu Phe Met Leu Gln Asp Ser Gly Ala Arg Leu Leu Val His Ser Gly
625 630 635 640

Pro Leu Pro Phe Ala Pro Gln Gly Cys Ala Thr Leu Asp Leu Gln Ala
645 650 655

Trp His Pro Ala Pro Ser Asp Gly Gly Glu Pro Leu Pro Gln Cys Ser
660 665 670

Gly Glu Asp Leu Ala Tyr Val Ile Tyr Thr Ser Gly Ser Thr Gly Gln
675 680 685

Pro Lys Gly Val Ala Val Cys His Arg Ala Met Thr Arg Leu Val Cys
690 695 700

Asn Thr Asp Tyr Val Gln Leu Gly Pro Glu Asp Arg Val Ala Gln Ala
705 710 715 720

Ser Asn Ala Ser Phe Asp Ala Ala Thr Phe Glu Val Trp Gly Ala Leu
725 730 735

Leu Asn Gly Ala Arg Leu Val Gly Leu Ala Thr Glu Glu Ala Ile Gln
740 745 750

Ala Arg Arg Leu Ala Glu Val Leu Arg Glu Gln Arg Ile Ser Val Leu
755 760 765

Phe Val Thr Thr Ala Leu Phe Asn His Val Ala Arg Glu Gln Pro Gln
770 775 780

Ala Phe Ser Thr Leu Arg Tyr Leu Leu Phe Gly Gly Glu Ala Val Asp
 785 790 795 800
 Ala Ser Ser Val Arg Arg Val Leu Lys Gln Gly Ala Pro Gly His Leu
 805 810 815
 Leu His Val Tyr Gly Pro Thr Glu Asn Thr Thr Phe Ser Thr Ala Trp
 820 825 830
 Arg Val Glu His Leu Ala Glu Gln Ala His Thr Val Pro Met Gly His
 835 840 845
 Pro Ile Ala Asn Ser Arg Leu His Val Leu Asp Glu Ala Leu Gln Pro
 850 855 860
 Val Pro Val Gly Ala Met Gly Glu Val Tyr Leu Gly Gly Asp Gly Leu
 865 870 875 880
 Ala Leu Gly Tyr Trp Arg His Pro Glu Ala Thr Ala Glu Arg Phe Val
 885 890 895
 Pro Asp Pro His Gly Leu Glu Pro Gly Gly Arg Leu Tyr Arg Thr Gly
 900 905 910
 Asp Leu Ala Arg Arg Gln Ala Asp Gly Ala Val Val Phe Ala Gly Arg
 915 920 925
 Val Asp Arg Gln Val Lys Leu Arg Gly Phe Arg Val Glu Pro Ala Glu
 930 935 940
 Ile Glu Ser His Leu Cys Glu His Ser Glu Val Ser Ala Ala Val Val
 945 950 955 960
 Glu Leu Arg Gly Glu Gly Ala Leu Arg Arg Leu Val Ala Tyr Val Val
 965 970 975
 Pro Arg Ala Gly Gly Arg Pro Gly Ala Glu Glu Leu Arg Thr Phe Leu
 980 985 990
 Arg Thr Arg Leu Pro Glu Tyr Met Leu Pro Ala Ser Phe Ser Leu Leu
 995 1000 1005
 Glu Ala Leu Pro Leu Thr Pro Asn Gly Lys Val Asp Arg Ser Ala
 1010 1015 1020
 Leu Pro Glu Ser Phe Glu Glu Ala Ser Arg Glu Gln Ala Pro Val
 1025 1030 1035

Val 1040	Pro	Pro Arg Gly Pro	Val 1045	Glu Ala Leu Leu	Val 1050	Asp Ile Trp
Arg 1055	Glu	Val Leu Gly Thr	Gln 1060	Arg Val Ser Val	His 1065	Asp Asp Phe
Phe 1070	Asp	Leu Gly Gly His	Ser 1075	Leu Leu Ala Thr	Arg 1080	Val Val Ser
Arg 1085	Leu	Arg Glu Ala Leu	Gln 1090	Val Glu Leu Pro	Leu 1095	Arg Thr Leu
Phe 1100	Glu	Ala Pro Gln Leu	Ser 1105	Ala Leu Ala Ala	Gln 1110	Val Glu Val
Leu 1115	Leu	Gly His Arg Gln	Leu 1120	Arg Pro Pro Pro	Leu 1125	Val Pro Ala
Val 1130	Arg	Pro Pro Glu Leu	Pro 1135	Leu Ser Phe Ala	Gln 1140	Gln Arg Leu
Trp 1145	Phe	Leu Gln Gln Leu	Ala 1150	Pro Gln Ser Thr	Ala 1155	Tyr Gln Ile
Leu 1160	Asp	Ala Trp His Val	Arg 1165	Gly Arg Val Asp	Val 1170	Gly Ala Leu
Glu 1175	Arg	Ala Leu Glu Gln	Leu 1180	Val Arg Arg His	Glu 1185	Ala Leu Arg
Thr 1190	Thr	Phe Glu Pro Gly	Gly 1195	Asp Gly Val Pro	Arg 1200	Gln Arg Ile
His 1205	Ala	Pro Ala Pro Val	Pro 1210	Leu Arg Gln Val	Asp 1215	Leu Arg Ser
His 1220	Gly	Ile Ala Ala Arg	Glu 1225	Glu Ala Leu Arg	Trp 1230	Met Arg Glu
Gln 1235	Ala	Leu Arg Pro Leu	Glu 1240	Leu Asp Lys Gly	Pro 1245	Leu Leu Arg
Val 1250	Ser	Leu Leu Arg Leu	Glu 1255	Glu Ala Gln Ser	Leu 1260	Leu Phe Leu
Glu Leu	His His Ile Val Gly	Asp Gly Trp Ser Leu	Ser Val Trp			

1265

1270

1275

Ser Arg Glu Leu Ser His Leu Tyr Glu Ala Ala Leu His Gly Ala
1280 1285 1290

Glu Pro Gly Leu Ala Pro Leu Pro Val Gln Tyr Ala Asp Phe Ala
1295 1300 1305

Leu Trp Gln Arg Gly Trp Leu Gln Gly Pro Val Leu Arg Glu Glu
1310 1315 1320

Leu Thr Trp Trp Arg Glu Arg Leu Ala Arg Leu Ala Pro Leu Arg
1325 1330 1335

Leu Pro Ala Asp His Ala Arg Pro Glu Val Gln Arg Phe Asn Gly
1340 1345 1350

Ala Thr Tyr Arg Phe Thr Leu Pro Gly Val Arg Val Gln Ala Leu
1355 1360 1365

Arg Arg Leu Gly His Glu His Gly Ala Thr Leu Phe Met Val Leu
1370 1375 1380

Leu Ala Gly Phe Asn Ala Leu Leu Ala Arg Tyr Thr Gly Gln Thr
1385 1390 1395

Asp Ile Ala Ile Gly Ala Pro Ile Ala Asn Arg Thr Arg Gly Glu
1400 1405 1410

Val Glu Gly Leu Ile Gly Phe Phe Val Asn Thr Leu Val Leu Arg
1415 1420 1425

Thr Arg Leu Glu Gly Asn Pro Ser Phe Leu Glu Leu Leu Arg Arg
1430 1435 1440

Val Arg Glu Thr Thr Leu Glu Ala Tyr Ala His Gln Glu Leu Pro
1445 1450 1455

Phe Glu Arg Leu Val Glu Glu Leu Gln Pro Glu Arg Gln Ala Asn
1460 1465 1470

Gln Asn Pro Leu Val Gln Val Leu Leu Ala Leu Gln Asn Ala Pro
1475 1480 1485

Arg Glu Pro Leu Arg Leu Ala Gly Leu Glu Ala Glu His Leu Glu
1490 1495 1500

Tyr	Leu	Val	Ala	Thr	Thr	Arg	Phe	Asp	Leu	Glu	Leu	His	Leu	Trp
1505						1510						1515		
Glu	Glu	Glu	Glu	Gly	Leu	Ser	Cys	Ile	Ala	Val	Tyr	Asp	Arg	Ala
1520						1525						1530		
Leu	Tyr	Gly	Ala	Gly	Thr	Val	Glu	Arg	Leu	Val	Gly	Ala	Trp	Cys
1535						1540						1545		
Thr	Leu	Leu	Glu	Gly	Val	Ala	Glu	Leu	Pro	Ala	Arg	Arg	Val	Ala
1550						1555						1560		
Glu	Leu	Pro	Leu	Val	Pro	Ala	Arg	Glu	Leu	Arg	Gln	Val	Pro	Pro
1565						1570						1575		
Pro	Ser	Pro	Ala	Ala	Ala	Ile	Glu	Thr	Ser	Ile	Gly	Ala	Arg	Phe
1580						1585						1590		
Ser	Glu	Val	Ala	Arg	Arg	Gln	Pro	Gly	Ala	Thr	Ala	Val	Thr	Gln
1595						1600						1605		
Gly	Gly	Arg	His	Leu	Thr	Tyr	Ala	Glu	Leu	Glu	Glu	Arg	Ser	Glu
1610						1615						1620		
Arg	Leu	Ala	Arg	Tyr	Leu	Ala	Trp	Leu	Gly	Val	Arg	Ala	Gly	Asp
1625						1630						1635		
Arg	Val	Gly	Leu	Ala	Thr	Glu	Arg	Thr	Leu	Glu	Arg	Ile	Ile	Ser
1640						1645						1650		
Leu	Leu	Gly	Ile	Leu	Lys	Ala	Gly	Ala	Ala	Tyr	Val	Pro	Leu	Asp
1655						1660						1665		
Val	Arg	Gln	Pro	Ala	Arg	Arg	Leu	Ser	Leu	Leu	Val	Gln	Ala	Ala
1670						1675						1680		
Gly	Val	Arg	Thr	Val	Ile	Ala	Glu	Glu	Gln	Ala	Arg	Thr	Val	Leu
1685						1690						1695		
Ser	Gly	Leu	Gly	Gln	Pro	Leu	Thr	Leu	Val	Asp	Ala	Ala	Gln	Glu
1700						1705						1710		
Pro	Ala	Ser	Ala	Gln	Gln	Val	Pro	Ala	Leu	Gly	Pro	Glu	Arg	Ser
1715						1720						1725		
Leu	Gly	Gly	Asp	Met	Leu	Ala	Tyr	Val	Leu	Phe	Thr	Ser	Gly	Ser
1730						1735						1740		

Thr	Gly	Glu	Pro	Lys	Gly	Val	Cys	Ile	Pro	His	Arg	Ala	Val	Leu
1745						1750					1755			
Arg	Leu	Ile	His	Glu	Pro	Ser	Tyr	Val	Gln	Leu	Ser	Pro	Arg	Glu
1760						1765					1770			
Val	Met	Leu	His	Tyr	Ala	Pro	Leu	Glu	Phe	Asp	Ala	Ser	Thr	Phe
1775						1780					1785			
Glu	Val	Trp	Gly	Ala	Leu	Leu	Asn	Gly	Ala	Arg	Leu	Val	Leu	Val
1790						1795					1800			
Pro	Pro	Glu	Gln	Gln	Ser	Leu	Glu	Ser	Leu	Gly	Gln	Glu	Leu	Ser
1805						1810					1815			
Thr	Gln	Gly	Val	Thr	Val	Leu	Trp	Leu	Thr	Ala	Gly	Leu	Phe	Arg
1820						1825					1830			
Leu	Met	Val	Glu	Glu	Gln	Leu	Lys	Ser	Leu	Arg	Gly	Val	Arg	Gln
1835						1840					1845			
Leu	Leu	Ala	Gly	Gly	Asp	Val	Leu	Pro	Met	Pro	Gln	Val	Arg	Arg
1850						1855					1860			
Leu	Arg	Glu	Ala	Leu	Pro	Glu	Cys	Gln	Leu	Ile	Asn	Gly	Tyr	Gly
1865						1870					1875			
Pro	Thr	Glu	Ser	Cys	Thr	Phe	Thr	Cys	Cys	His	Arg	Val	Gly	Ser
1880						1885					1890			
Pro	Gln	Glu	Leu	Gly	Gly	Ser	Val	Pro	Ile	Gly	Thr	Pro	Ile	Asp
1895						1900					1905			
Leu	Gly	Trp	Val	Ser	Val	Val	Asp	Glu	Arg	Leu	Gln	Pro	Val	Pro
1910						1915					1920			
Asp	Gly	Ala	Pro	Gly	Glu	Leu	Leu	Val	Gly	Gly	Pro	Gly	Leu	Ala
1925						1930					1935			
Trp	Gly	Tyr	Leu	Gln	His	Pro	Glu	Leu	Thr	Ala	Glu	Arg	Phe	Ile
1940						1945					1950			
Pro	Asp	Pro	Leu	Ser	Arg	Thr	Pro	Gly	Ala	Arg	Val	Tyr	Arg	Thr
1955						1960					1965			
Gly	Asp	Leu	Val	Arg	Arg	Arg	Glu	Asp	Gly	Thr	Leu	Glu	Phe	Leu
1970						1975					1980			

Gly	Arg	Val	Asp	His	Gln	Leu	Lys	Val	Arg	Gly	Phe	Arg	Ile	Glu
1985						1990					1995			
Pro	Gly	Glu	Val	Glu	Ala	Ala	Val	Leu	Thr	His	Pro	Ala	Val	Gln
2000						2005					2010			
Ser	Ala	Val	Val	Val	Gly	Arg	Glu	Gly	Pro	Gly	Gly	Lys	Glu	Leu
2015						2020					2025			
Val	Cys	Tyr	Ala	Val	Pro	Arg	Val	Glu	Ser	Ser	Glu	Gln	Gly	Ser
2030						2035					2040			
Gln	Gln	Glu	Gln	Arg	Leu	Val	His	Glu	Trp	Glu	Ser	Val	Phe	Asp
2045						2050					2055			
Gly	His	Met	Tyr	Arg	Glu	Ala	Pro	Val	Gly	Gly	Glu	Pro	Thr	Phe
2060						2065					2070			
Asn	Ile	Val	Gly	Trp	Lys	Ser	Ser	Tyr	Thr	Gly	Gln	Pro	Val	Ala
2075						2080					2085			
Val	Glu	Glu	Met	Arg	Asp	Trp	Leu	Arg	His	Arg	Val	Glu	Arg	Val
2090						2095					2100			
Arg	Gly	Leu	Arg	Pro	Arg	Arg	Ile	Leu	Glu	Val	Gly	Cys	Gly	Thr
2105						2110					2115			
Gly	Leu	Met	Leu	Phe	Ala	Leu	Leu	Pro	His	Cys	Glu	Arg	Tyr	Val
2120						2125					2130			
Gly	Thr	Asp	Phe	Ser	Pro	Ala	Ala	Leu	Asp	Tyr	Val	Arg	Arg	Tyr
2135						2140					2145			
Leu	Pro	Pro	Glu	His	Pro	Gly	Arg	Val	Glu	Leu	Leu	His	Arg	Thr
2150						2155					2160			
Ala	Asp	Glu	Trp	Ser	Gly	Val	Ala	Ala	Gly	Ser	Phe	Asp	Ala	Val
2165						2170					2175			
Leu	Leu	Asn	Ser	Val	Val	Gln	Tyr	Phe	Pro	Ser	Gln	Glu	Tyr	Leu
2180						2185					2190			
Arg	Gln	Val	Leu	Ala	Arg	Cys	Val	Glu	Ala	Val	Glu	Asp	Gly	Gly
2195						2200					2205			
Phe	Val	Phe	Val	Gly	Asp	Val	Arg	Ser	Leu	Pro	Leu	Leu	Glu	Ser

2210

2215

2220

Phe His Ala Ser Val Glu Leu Glu Arg Ala Ala Pro Ser Met Pro
2225 2230 2235

Leu Glu Ala Trp Arg Glu Arg Val Arg Arg Ala Val Leu Glu Asp
2240 2245 2250

Asn Glu Leu Val Val Asp Pro Ala Leu Phe Val Ala Leu Ala His
2255 2260 2265

Gln His Pro Arg Val Ser His Val Asp Ile Glu Leu Thr Arg Gly
2270 2275 2280

Thr His Pro Asn Glu Met Ala Arg Phe Arg Tyr Asn Ala Val Leu
2285 2290 2295

His Ile Gly Pro Arg Thr Pro Pro Pro Ala Ser Glu Val Pro Trp
2300 2305 2310

Val Asp Trp Ser Thr Gln Gly Leu Ser Leu Asp Ala Leu Arg Ala
2315 2320 2325

Arg Leu Arg Gln Gly Pro Pro Gly Pro Leu Gly Val Ala Gly Ile
2330 2335 2340

Pro Asn Ala Arg Val Leu Pro Ala Val Arg Ala Ala Glu Ala Leu
2345 2350 2355

Gly Ser Thr Gly Ser Ala Arg Arg Val Glu Glu Leu Arg Arg Arg
2360 2365 2370

Leu Ser Gln Pro Gly Thr Gly Ala Gln Asp Pro Asp Pro Phe Trp
2375 2380 2385

Gln Leu Ala Glu Ser Leu Gly Tyr Thr Ala Ala Val Ser Trp Ser
2390 2395 2400

Pro Gly Arg Arg Asp Gly Ala Phe Asp Val Leu Phe Leu Pro Ala
2405 2410 2415

Thr Pro Gly Met His Pro Arg Trp Leu Gly Pro Thr Pro Leu Asn
2420 2425 2430

Asp Thr Pro Pro Pro Ala Ser Ala Arg Leu Ser Ser Glu Pro Arg
2435 2440 2445

Arg Ala Ser Leu Ser Leu Arg Leu Gly Ser Ala Leu Arg Ala His
 2450 2455 2460

Leu Gln Thr His Leu Pro Asp Phe Met Val Pro Ser Arg Phe Val
 2465 2470 2475

Val Leu Gln Ser Leu Pro Leu Thr Pro Asn Gly Lys Val Asp Arg
 2480 2485 2490

Ala Ala Leu Pro Val Pro Asp Ser Arg Arg Leu Glu Ser Ala Pro
 2495 2500 2505

Leu Val Pro Pro Ser Asn Glu Leu Glu Arg Val Leu Ala Gln Val
 2510 2515 2520

Trp Lys Glu Val Leu Gly Leu Glu Glu Val Ser Arg Glu Asp Asn
 2525 2530 2535

Phe Phe Asp Val Gly Gly His Ser Leu Leu Leu Ala Gln Val Cys
 2540 2545 2550

Ser Arg Leu Glu Ala Arg Leu Gly Arg Arg Leu Glu Leu Val Thr
 2555 2560 2565

Leu Phe Arg Tyr Ser Ser Ile Ala Ala Leu Ala Glu His Leu Gln
 2570 2575 2580

Ala Pro Gln Glu Leu Ala Ala Ala Glu Ala Gln Val Gln Arg Met
 2585 2590 2595

Ala Thr Glu Arg Ala Leu Leu Gln Gln Gln Ala Ala Gln Arg Arg
 2600 2605 2610

Arg Ala Gly Ser Lys Arg Gly Ala Pro Asn Asp Thr
 2615 2620 2625

<210> 35

<211> 1181

<212> PRT

<213> *Angiococcus disciformis*

<400> 35

Met Ser Lys Leu His Glu Glu Leu Glu Ser Leu Ala Pro Glu Gln Arg
 1 5 10 15

Glu Leu Leu Ala Ala Leu Met Lys Glu Gln Gly Leu Asp Glu Gly Ala
20 25 30

Leu Leu Met Pro Val Glu Arg Lys Pro Glu Gly Leu Pro Leu Ser Ser
35 40 45

Ala Gln Gln Arg Met Trp Phe Leu Gln Gln Leu Arg Pro Thr Ser Pro
50 55 60

Phe Tyr Asn Val His Ala Ala Leu Arg Leu Thr Gly Lys Leu Lys Val
65 70 75 80

Glu Cys Leu Val His Ser Leu Asn Glu Phe Val Arg Arg His Glu Pro
85 90 95

Leu Arg Thr Val Phe Pro Ser Ala Gly Gly Gln Pro Leu Gln Arg Ile
100 105 110

Leu Ala Pro Ala Pro Ala Ala Leu Glu Gln Arg Asp Leu Ser Gly Val
115 120 125

Pro Ala Gln Glu Arg Glu Ala Glu Val Tyr Arg Ala Val Glu His Ala
130 135 140

Ala Leu Ala Ser Phe Asp Leu Glu Arg Glu Pro Pro Cys Arg Phe Leu
145 150 155 160

Leu Val Arg Val Glu Pro Glu Glu His Val Leu Val Phe Ala Thr His
165 170 175

His Ile Ala Ala Asp Gly Trp Ser Leu Gly Val Phe Val Gln Glu Leu
180 185 190

Cys Ala Leu Tyr Thr Ala Ala Val His Ala Glu Pro Pro Ala Leu Pro
195 200 205

Pro Leu Arg Leu Gln Tyr Ala Asp Phe Ala Ala Trp Glu Arg Ser Arg
210 215 220

Leu Lys Gly Gly Arg Glu Arg Glu Leu Leu Glu Tyr Trp Gln Glu Gln
225 230 235 240

Leu Ala Gly Leu Pro Asp Leu Ser Thr Leu Pro Pro Glu Arg Pro Arg
245 250 255

Pro Pro Leu Ser Lys Gln Glu Gly Ala Thr Phe Glu Phe Ala Leu Pro
260 265 270

Pro His Gln Val Gln Ala Leu Arg Ser Leu Ala Gln Ala Arg Arg Thr
275 280 285

Ser Leu Phe Ser Val Leu Leu Ala Ala Phe Gln Trp Leu Leu Ala Arg
290 295 300

Cys Ala Gly Gln Asp Asp Val Ala Leu Gly Met Pro Ile Ala Asn Arg
305 310 315 320

Asn Arg Lys Glu Leu Glu Gly Leu Ile Gly Cys Phe Ala Ser Thr Leu
325 330 335

Val Leu Arg Ala Lys Pro Ser Ala Ser Thr Ala Phe Thr Ser Trp Leu
340 345 350

Ala Gln Val Ser Glu Gln Leu His Gly Ala Leu Glu His Gln Glu Val
355 360 365

Pro Phe Glu Arg Leu Val Glu Val Leu Gln Pro Arg Arg Arg Met Asp
370 375 380

Arg His Pro Leu Phe Gln Ile Phe Leu Ala Met Gln Gln His Pro Leu
385 390 395 400

Arg Arg Ala Glu Leu Pro Gly Leu Leu Leu Ser Glu Phe Pro Leu Arg
405 410 415

Ser Arg Val Ala Arg Phe Asp Leu Glu Phe His Leu Trp Glu Ser Ala
420 425 430

Gln Gly Val Glu Gly Thr Leu Ile Tyr Asp Val Asp Leu Tyr Gly Gln
435 440 445

Glu Gly Val Ala Gln Leu Val Arg Arg Tyr Val Ser Leu Leu Glu Ala
450 455 460

Val Ala Ala Asp Pro Thr Arg Thr Leu Gly Glu Leu Ala Gly Glu Thr
465 470 475 480

Leu Ala Pro Pro Val Arg Lys Gln Val Leu Ala Leu Ser Glu Cys Pro
485 490 495

Pro Leu Pro Arg Pro Ser Thr Ala Pro Arg Thr Leu Ala Glu Ala Leu
500 505 510

Leu Gln Thr Ala Glu Arg Phe Pro Thr Ala Thr Val Ser Phe Val Gln

515

520

525

Ala Gln Gly Ser Cys Thr Ala Trp Thr Leu Pro Glu Leu Val Glu Arg
530 535 540

Ala Arg Arg Leu Gln Ala Gly Leu Arg Gln Trp Gly Leu Arg Pro Gly
545 550 555 560

Asp Ser Leu Val Leu Val Leu Gly Arg Glu Glu Glu Thr Val Glu Ala
565 570 575

Leu Trp Ala Cys Val Leu Ala Gly Val Ala Pro Leu Val Leu Pro Ala
580 585 590

Pro Pro Ala Arg Ala Glu Ala Ser Pro Ala Leu Ser Arg Leu Arg His
595 600 605

Ala Arg Gln Leu Leu Gly Gly Pro Arg Val Leu Thr Arg Gln Glu Met
610 615 620

Leu Pro Asp Leu Ala Arg Gln Leu Gln Val Ser Pro Thr Ala Asp Ile
625 630 635 640

Leu Gly Ala Val Glu Glu Leu Arg Ala Thr Gly Gly Glu Ala Pro Leu
645 650 655

Pro Pro Gly Arg Met Asp Asp Val Ala Leu Leu Asn Leu Thr Ser Gly
660 665 670

Thr Thr Gly Lys Ala Lys Cys Ala Met Leu Thr His Arg Asn Leu Leu
675 680 685

Val Arg Leu Glu Ala Thr Asn Val Val Tyr Glu Ser Gln Pro Leu Glu
690 695 700

Arg Gly Leu Val Trp Leu Gln Leu His Asn Ile Gly Ala Leu Ser Glu
705 710 715 720

Tyr His Leu Arg Pro Leu Cys Ala Gly Met His Thr Phe His Ala Pro
725 730 735

Thr Glu Glu Val Leu Ala Glu Pro Leu Arg Trp Leu Glu Trp Leu Glu
740 745 750

Arg Tyr Gly Ile Ala Gln Thr Trp Ala Pro Ser Phe Ala Tyr Ser His
755 760 765

Leu Leu Glu Arg Leu Arg Lys Val Glu Asp Arg Arg Trp Ser Leu Gly
770 775 780

Gly Val Arg Val Leu Leu Ser Ala Gly Glu Gln Ile Ser Ala Pro Met
785 790 795 800

Val Glu Glu Leu Met Arg Arg Leu Ala Pro Ser Gly Val Arg Glu Asp
805 810 815

Ala Phe Val Ala Ala Trp Gly Met Thr Glu Thr Ala Ser Gly Val Thr
820 825 830

Tyr Ala Arg Arg Pro Gly Thr Pro Pro Arg Met His Thr Leu Glu Arg
835 840 845

Ala Ser Leu Ser Gly Pro Leu Arg His Ala Ala Pro Ala Ser Pro Thr
850 855 860

Ala Leu Arg Leu Met Asp Val Gly Ala Pro Ile Ala Gly Thr Ala Leu
865 870 875 880

Arg Val Val Asp Ala Ser Gly Glu Leu Leu Ser Glu Glu Cys Val Gly
885 890 895

Arg Ile Gln Val Arg Gly Glu Met Ile Ser Pro Gly Tyr Tyr Gly Asp
900 905 910

Pro Lys Ala Ser Ala Ala Leu Leu Thr Ala Asp Gly Trp Leu Glu Thr
915 920 925

Gly Asp Leu Gly Phe Leu Ser Glu Gly Ala Leu Thr Ile Thr Gly Arg
930 935 940

Ala Lys Asp Leu Val Ile Ile His Gly Thr Asn Phe Ser Cys Tyr Glu
945 950 955 960

Ile Glu Ser Ala Val Glu Gln Val Glu Gly Val Ala Pro Ser Ser Ala
965 970 975

Ala Ala Ala Ala Val Arg Met Leu Glu Gly Ser Arg Glu Glu Leu Ala
980 985 990

Val Phe Phe Val Pro Thr Glu Gly Leu Ala Pro Gln Pro Pro Ala Ser
995 1000 1005

Leu Leu Ser Arg Ile Arg Gln Gln Val Leu Glu Gln Val Gly Val
1010 1015 1020

Arg Ile Asp His Leu Ile Pro Leu Glu Pro His Gln Leu Pro Arg
1025 1030 1035

Thr Glu Gly Gly Lys Leu Arg Arg Ser Glu Leu Arg Ala Arg Phe
1040 1045 1050

Glu Ala Gly Glu Leu Arg Ala Pro Gln Pro Ala Pro Val Pro Ser
1055 1060 1065

Pro Ser Arg Pro Leu Glu Gln Leu Ile Ala Ser Val Trp Ala Glu
1070 1075 1080

Val Leu Glu His Gln Asp Ile Ala Pro Glu Ala Ser Phe Phe Asp
1085 1090 1095

Leu Gly Gly Asn Ser Ile Leu Leu Val Arg Val Glu Arg Ala Leu
1100 1105 1110

Arg Ala Arg Leu Gly Leu Glu Leu Thr Leu Met Asp Leu Phe Ala
1115 1120 1125

Tyr Pro Thr Val His Ser Leu Ala Asp Tyr Leu Glu Pro Arg Ala
1130 1135 1140

Ala Gln Leu Pro Ala Gln Ala Ser Pro Pro Thr Gln Ala Glu Arg
1145 1150 1155

Arg Arg Gly Met Arg Gly Ala Ala Leu Glu Gln Arg Arg Thr Arg
1160 1165 1170

Arg Gln Ala Gln Arg Asp Glu Asp
1175 1180

<210> 36

<211> 2200

<212> DNA

<213> *Angiococcus disciformis*

<220>

<221> CDS

<222> (44) .. (970)

<223> ORF26

<220>

<221> CDS

<222> (973)..(2181)

<223> ORF25

<220>

<221> misc_feature

<222> (2195)..(2195)

<223> n=A, C, G or T

<400> 36

```

gggagcgcgc tcggcaaggg cagggcgtctg ttaggggtcta cct atg agg tta cgt      55
                                   Met Arg Leu Arg
                                   1

ccc atg ccc cac gtt ccc ggg cta ttg ctg gtg ctg ctg tcc atg gcc      103
Pro Met Pro His Val Pro Gly Leu Leu Leu Val Leu Leu Ser Met Ala
5                               10                               15                               20

gca tgg gca gcc gaa gcg cag acg gcc ccc ggc cct tcc ctt cgt ttc      151
Ala Trp Ala Ala Glu Ala Gln Thr Ala Pro Gly Pro Ser Leu Arg Phe
                               25                               30                               35

cgc cgc gtg ccc ctc act gga aag ccc gag gtg gtg cgg ctc gca acg      199
Arg Arg Val Pro Leu Thr Gly Lys Pro Glu Val Val Arg Leu Ala Thr
                               40                               45                               50

ggc att ccc gcg acg ctc aac ttc gac gcg gag atc aac ccg cag aaa      247
Gly Ile Pro Ala Thr Leu Asn Phe Asp Ala Glu Ile Asn Pro Gln Lys
                               55                               60                               65

gtg cgg ctg gat gct ccg gcc cgc gtc aag att ttg agc acc ggg gag      295
Val Arg Leu Asp Ala Pro Ala Arg Val Lys Ile Leu Ser Thr Gly Glu
                               70                               75                               80

cgc tcc gtc acc ctg ctt gcc gag ggc gat ttg ggc gaa ggg gtg act      343
Arg Ser Val Thr Leu Leu Ala Glu Gly Asp Leu Gly Glu Gly Val Thr
85                               90                               95                               100

ttg cgc gtg ccg ttc ctg act gag ccc tcc ctc gcg gaa ccc gcc ttt      391
Leu Arg Val Pro Phe Leu Thr Glu Pro Ser Leu Ala Glu Pro Ala Phe
                               105                               110                               115

cag ctc gtc acg gct tcg gac gtg gcg gat tcg cag gtg atg gtc tac      439
Gln Leu Val Thr Ala Ser Asp Val Ala Asp Ser Gln Val Met Val Tyr
                               120                               125                               130

cgc agc gcc aac gcg ccg gag ttg atg caa gca cgg ctg acg gag ttg      487
Arg Ser Ala Asn Ala Pro Glu Leu Met Gln Ala Arg Leu Thr Glu Leu
                               135                               140                               145

```

gag gcc cgc agc gcc gcg tgc gag gcg caa ctt tcc acc gag cgc gag Glu Ala Arg Ser Ala Ala Cys Glu Ala Gln Leu Ser Thr Glu Arg Glu 150 155 160	535
cgc agc gag gcc acg ggc ccc gcc gcc tgg gtc ctc tcc cgg caa gtt Arg Ser Glu Ala Thr Gly Pro Ala Ala Trp Val Leu Ser Arg Gln Val 165 170 175 180	583
gat gag ggc ggc gtg agc gtt gtc atg ctc aag cga ctg tcc gac gtg Asp Glu Gly Gly Val Ser Val Val Met Leu Lys Arg Leu Ser Asp Val 185 190 195	631
ggc acg gcg gcc att gag gca acg tcc gtt cgc cgc ttc ctg gcg gat Gly Thr Ala Gly Ile Glu Ala Thr Ser Val Arg Arg Phe Leu Ala Asp 200 205 210	679
aag tgg gcg gtg ctc caa gtg caa gtg agg aac ggc tct ggg cag cct Lys Trp Ala Val Leu Gln Val Gln Val Arg Asn Gly Ser Gly Gln Pro 215 220 225	727
tgg cgg cca ggg cgg gca tgg ctg gaa agt gcg gcc acc agg cgc ctc Trp Arg Pro Gly Arg Ala Trp Leu Glu Ser Ala Ala Thr Arg Arg Leu 230 235 240	775
acg gaa gcg cgc acg gtg agc atg gcg ccg gaa gtg ctg gcg ccc gaa Thr Glu Ala Arg Thr Val Ser Met Ala Pro Glu Val Leu Ala Pro Glu 245 250 255 260	823
ggt aag agc cgc gtg gtg gtg gag ttt gcc ggg aag ccc ggg agc aag Gly Lys Ser Arg Val Val Val Glu Phe Ala Gly Lys Pro Gly Ser Lys 265 270 275	871
ggg gaa gtc ttc cga ttg gtg gta gag gac gca gac ggc gtc cgg ccc Gly Glu Val Phe Arg Leu Val Val Glu Asp Ala Asp Gly Val Arg Pro 280 285 290	919
ctc tca gtc gcg ggc gtg gtg atg aag gac aca gcg cag gag aag agc Leu Ser Val Ala Gly Val Val Met Lys Asp Thr Ala Gln Glu Lys Ser 295 300 305	967
ggc cc atg atg acc gtg gat ccc ccc aac ttc ctc aag ccc ggc gcc Gly Met Met Thr Val Asp Pro Pro Asn Phe Leu Lys Pro Gly Ala 310 315 320	1014
cgc ctc ggg gcc ccc tac cgg att cgc ggt tgg aag ggc ggc ggt gcc Arg Leu Gly Ala Pro Tyr Arg Ile Arg Gly Trp Lys Gly Gly Gly Ala 325 330 335	1062
tat ggc gag gtg tac gag tgc gaa agg cag ggc agg ccc tac gcc ctc Tyr Gly Glu Val Tyr Glu Cys Glu Arg Gln Gly Arg Pro Tyr Ala Leu 340 345 350 355	1110
aag ctg tcc aag cac cgc cag tcc agc gat gac ccg ggc aag acg gac Lys Leu Ser Lys His Arg Gln Ser Ser Asp Asp Pro Gly Lys Thr Asp 360 365 370	1158
caa cgg ctg ctg cgg gag ttg gtg tgc ctc gtc caa ctg gac cac ccg Gln Arg Leu Leu Arg Glu Leu Val Cys Leu Val Gln Leu Asp His Pro 375 380 385	1206
aac att gcc aag gtg ctg ggc tgg aca cgc act gac gag ggt tac ggc Asn Ile Ala Lys Val Leu Gly Trp Thr Arg Thr Asp Glu Gly Tyr Gly 390 395 400	1254

tat ctg gtg ctg gag tac atc gag ggg tgg acg ctg gcc caa tgg ctc Tyr Leu Val Leu Glu Tyr Ile Glu Gly Trp Thr Leu Ala Gln Trp Leu 405 410 415	1302
cag caa acg cgc ccc acc gtc gcc caa gtc ctg cac gtc ttc gcc aaa Gln Gln Thr Arg Pro Thr Val Ala Gln Val Leu His Val Phe Ala Lys 420 425 430 435	1350
ctg acg ggg gcg ctc gcg tac atg cac cgc agg ggc gtt cgg cac cgg Leu Thr Gly Ala Leu Ala Tyr Met His Arg Arg Gly Val Arg His Arg 440 445 450	1398
gac gtg tgc ctg tct aac gtc atg gta cgc aag gcg gac ggc gag ccc Asp Val Ser Leu Ser Asn Val Met Val Arg Lys Ala Asp Gly Glu Pro 455 460 465	1446
gtc ctc att gac ctc ggg gcg ggg gac tac tcc ggg gcg cat gag ttg Val Leu Ile Asp Leu Gly Ala Gly Asp Tyr Ser Gly Ala His Glu Leu 470 475 480	1494
acg gac tgg ccc ttg ccc ccg ggc acc gac cgc tac cgc tck ccc gar Thr Asp Trp Pro Leu Pro Pro Gly Thr Asp Arg Tyr Arg Xaa Pro Glu 485 490 495	1542
gcg gcc cgc ttc ctc aag gya cac cag aac gac ccg agc gcc cgg tac Ala Ala Arg Phe Leu Lys Xaa His Gln Asn Asp Pro Ser Ala Arg Tyr 500 505 510 515	1590
gac ttc ccc ccg gaa gat gac ttc tac tgc ctg gcg gtg tgc ctc tat Asp Phe Pro Pro Glu Asp Asp Phe Tyr Ser Leu Ala Val Cys Leu Tyr 520 525 530	1638
gac gcg ctg acg gac gcg gag ccc gcc ctc aag gcg aac gca cgc aag Asp Ala Leu Thr Asp Ala Glu Pro Ala Leu Lys Ala Asn Ala Arg Lys 535 540 545	1686
gcc ccg cgc ctc aac gtc aac agc ccc acc atg gcc ccg ccc ccg gcc Ala Pro Arg Leu Asn Val Asn Ser Pro Thr Met Ala Pro Pro Pro Ala 550 555 560	1734
cgc aag gtc aac ccg cgc gtg ccc gag gta gtg agt gcc tgg gta gag Arg Lys Val Asn Pro Arg Val Pro Glu Val Val Ser Ala Trp Val Glu 565 570 575	1782
cgg tgg ctg gtg cgt gac gtg gaa acc cgg cgg ccc gcg ttg gcc gcc Arg Trp Leu Val Arg Asp Val Glu Thr Arg Arg Pro Ala Leu Ala Ala 580 585 590 595	1830
atg ccc ggc gcg ctg gaa gag ttg gcg aag caa ggc ggc gcg gag tgg Met Pro Gly Ala Leu Glu Glu Leu Ala Lys Gln Gly Gly Ala Glu Trp 600 605 610	1878
ctg gcc acc gtt cag ccc cca ccg gaa gcg ggc acc ccg gcc ccc gag Leu Ala Thr Val Gln Pro Pro Pro Glu Ala Gly Thr Pro Ala Pro Glu 615 620 625	1926
cac cgg ctc tac cgg aag cgg cgg cgc gtg ctg gcg tgg gcc atg gcc His Arg Leu Tyr Arg Lys Arg Arg Arg Val Leu Ala Trp Ala Met Ala 630 635 640	1974
gct ggc gtg ctc gtg gct gtc gcg gcg ctc gcg tgg ctg cgc gct gcc Ala Gly Val Leu Val Ala Val Ala Ala Leu Ala Trp Leu Arg Ala Ala	2022

645

650

655

ccc tct cca gac gcg ccg ccg tcc cag gtc gta ccg gag gca cca agc 2070
 Pro Ser Pro Asp Ala Pro Pro Ser Gln Val Val Pro Glu Ala Pro Ser
 660 665 670 675

acg ccg cag gcg cca gcc ccg gac agt cca cca acc ccg ccc gtg ccg 2118
 Thr Pro Gln Ala Pro Ala Pro Asp Ser Pro Pro Thr Pro Pro Val Pro
 680 685 690

tct tcg ggt gca gta ccg gtc ccc atc cca ggt gtc cct gtt gag aag 2166
 Ser Ser Gly Ala Val Pro Val Pro Ile Pro Gly Val Pro Val Glu Lys
 695 700 705

gaa agc cct ctg tga gtacccccac gaantccgg 2200
 Glu Ser Pro Leu
 710

<210> 37

<211> 309

<212> PRT

<213> *Angiococcus disciformis*

<220>

<221> misc_feature

<222> (2195)..(2195)

<223> n=A, C, G or T

<400> 37

Met Arg Leu Arg Pro Met Pro His Val Pro Gly Leu Leu Leu Val Leu
 1 5 10 15

Leu Ser Met Ala Ala Trp Ala Ala Glu Ala Gln Thr Ala Pro Gly Pro
 20 25 30

Ser Leu Arg Phe Arg Arg Val Pro Leu Thr Gly Lys Pro Glu Val Val
 35 40 45

Arg Leu Ala Thr Gly Ile Pro Ala Thr Leu Asn Phe Asp Ala Glu Ile
 50 55 60

Asn Pro Gln Lys Val Arg Leu Asp Ala Pro Ala Arg Val Lys Ile Leu
 65 70 75 80

Ser Thr Gly Glu Arg Ser Val Thr Leu Leu Ala Glu Gly Asp Leu Gly
 85 90 95

Glu Gly Val Thr Leu Arg Val Pro Phe Leu Thr Glu Pro Ser Leu Ala
100 105 110

Glu Pro Ala Phe Gln Leu Val Thr Ala Ser Asp Val Ala Asp Ser Gln
115 120 125

Val Met Val Tyr Arg Ser Ala Asn Ala Pro Glu Leu Met Gln Ala Arg
130 135 140

Leu Thr Glu Leu Glu Ala Arg Ser Ala Ala Cys Glu Ala Gln Leu Ser
145 150 155 160

Thr Glu Arg Glu Arg Ser Glu Ala Thr Gly Pro Ala Ala Trp Val Leu
165 170 175

Ser Arg Gln Val Asp Glu Gly Gly Val Ser Val Val Met Leu Lys Arg
180 185 190

Leu Ser Asp Val Gly Thr Ala Gly Ile Glu Ala Thr Ser Val Arg Arg
195 200 205

Phe Leu Ala Asp Lys Trp Ala Val Leu Gln Val Gln Val Arg Asn Gly
210 215 220

Ser Gly Gln Pro Trp Arg Pro Gly Arg Ala Trp Leu Glu Ser Ala Ala
225 230 235 240

Thr Arg Arg Leu Thr Glu Ala Arg Thr Val Ser Met Ala Pro Glu Val
245 250 255

Leu Ala Pro Glu Gly Lys Ser Arg Val Val Val Glu Phe Ala Gly Lys
260 265 270

Pro Gly Ser Lys Gly Glu Val Phe Arg Leu Val Val Glu Asp Ala Asp
275 280 285

Gly Val Arg Pro Leu Ser Val Ala Gly Val Val Met Lys Asp Thr Ala
290 295 300

Gln Glu Lys Ser Gly
305

<210> 38

<211> 402

<212> PRT

<213> Angiococcus disciformis

<220>

<221> misc_feature

<222> (188)..(188)

<223> The 'Xaa' at location 188 stands for Ser.

<220>

<221> misc_feature

<222> (197)..(197)

<223> The 'Xaa' at location 197 stands for Ala, or Val.

<220>

<221> misc_feature

<222> (2195)..(2195)

<223> n=A, C, G or T

<400> 38

Met Met Thr Val Asp Pro Pro Asn Phe Leu Lys Pro Gly Ala Arg Leu
1 5 10 15

Gly Ala Pro Tyr Arg Ile Arg Gly Trp Lys Gly Gly Gly Ala Tyr Gly
20 25 30

Glu Val Tyr Glu Cys Glu Arg Gln Gly Arg Pro Tyr Ala Leu Lys Leu
35 40 45

Ser Lys His Arg Gln Ser Ser Asp Asp Pro Gly Lys Thr Asp Gln Arg
50 55 60

Leu Leu Arg Glu Leu Val Cys Leu Val Gln Leu Asp His Pro Asn Ile
65 70 75 80

Ala Lys Val Leu Gly Trp Thr Arg Thr Asp Glu Gly Tyr Gly Tyr Leu
85 90 95

Val Leu Glu Tyr Ile Glu Gly Trp Thr Leu Ala Gln Trp Leu Gln Gln
100 105 110

Met Arg Pro Thr Val Ala Gln Val Leu His Val Phe Ala Lys Leu Thr
115 120 125

Gly Ala Leu Ala Tyr Met His Arg Arg Gly Val Arg His Arg Asp Val

130

135

140

Ser Leu Ser Asn Val Met Val Arg Lys Ala Asp Gly Glu Pro Val Leu
145 150 155 160

Ile Asp Leu Gly Ala Gly Asp Tyr Ser Gly Ala His Glu Leu Thr Asp
165 170 175

Trp Pro Leu Pro Pro Gly Thr Asp Arg Tyr Arg Xaa Pro Glu Ala Ala
180 185 190

Arg Phe Leu Lys Xaa His Gln Asn Asp Pro Ser Ala Arg Tyr Asp Phe
195 200 205

Pro Pro Glu Asp Asp Phe Tyr Ser Leu Ala Val Cys Leu Tyr Asp Ala
210 215 220

Leu Thr Asp Ala Glu Pro Ala Leu Lys Ala Asn Ala Arg Lys Ala Pro
225 230 235 240

Arg Leu Asn Val Asn Ser Pro Thr Met Ala Pro Pro Pro Ala Arg Lys
245 250 255

Val Asn Pro Arg Val Pro Glu Val Val Ser Ala Trp Val Glu Arg Trp
260 265 270

Leu Val Arg Asp Val Glu Thr Arg Arg Pro Ala Leu Ala Ala Met Pro
275 280 285

Gly Ala Leu Glu Glu Leu Ala Lys Gln Gly Gly Ala Glu Trp Leu Ala
290 295 300

Thr Val Gln Pro Pro Pro Glu Ala Gly Thr Pro Ala Pro Glu His Arg
305 310 315 320

Leu Tyr Arg Lys Arg Arg Arg Val Leu Ala Trp Ala Met Ala Ala Gly
325 330 335

Val Leu Val Ala Val Ala Ala Leu Ala Trp Leu Arg Ala Ala Pro Ser
340 345 350

Pro Asp Ala Pro Pro Ser Gln Val Val Pro Glu Ala Pro Ser Thr Pro
355 360 365

Gln Ala Pro Ala Pro Asp Ser Pro Pro Thr Pro Pro Val Pro Ser Ser
370 375 380

Gly Ala Val Pro Val Pro Ile Pro Gly Val Pro Val Glu Lys Glu Ser
385 390 395 400

Pro Leu

<210> 39

<211> 21

<212> DNA

<213> artificial sequence

<220>

<223> K-388 primer

<400> 39

tgggaatcat ttgaagggtg g

21

<210> 40

<211> 22

<212> DNA

<213> artificial sequence

<220>

<223> K-389 primer

<400> 40

tgtgtttttc tttgttagac cg

22

<210> 41

<211> 21

<212> DNA

<213> artificial sequence

<220>

<223> ASTIs1A primer

<400> 41

cacccggacc tgcctggatt c

21

<210> 42

<211> 21

<212> DNA

<213> artificial sequence

<220>

<223> ASTIs1B primer

<400> 42

tgctcggctg gcgctactca c

21

<210> 43

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> ASTIs2A primer

<400> 43

gctcccgggc cacgtggttg aaga

24

<210> 44

<211> 23

<212> DNA

<213> artificial sequence

<220>

<223> ASTIs2B primer

<400> 44

ccgcgggccg tggcagtggg gta

23

<210> 45

<211> 17

<212> DNA

<213> artificial sequence

<220>

<223> TIs-up primer

<400> 45

tggcagccag cccgagc

17

<210> 46

<211> 19

<212> DNA

<213> artificial sequence

<220>

<223> TIs-down primer

<400> 46

ccgcgggtgc cctctcatc

19